

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 11.5063 Seconds  
(without alignments)  
731.761 Million cell updates/sec

Title: US-10-083-336A-5  
Perfect score: 1025  
Sequence: 1 MIPPKYPIINFITAGATVQ.....ARQYIEGEMRTIRYNRRS 199

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pcp.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1025	100.0	268	2	US-08-356-786-8
2	1025	100.0	534	2	US-08-356-786-10
3	1020	99.5	267	1	US-07-901-707-1
4	1020	99.5	267	1	US-07-988-430-1
5	1020	99.5	267	1	US-08-425-336-1
6	1020	99.5	267	1	US-08-488-113B-1
7	1020	99.5	267	1	US-08-477-484B-1
8	1020	99.5	267	2	US-08-646-360-1
9	1020	99.5	267	3	US-08-839-765-1
10	1020	99.5	267	3	US-09-136-389-1
11	1020	99.5	267	4	US-09-610-838-1
12	1020	99.5	267	5	PCT-US92-09487-1
13	1020	99.5	290	1	US-08-378-761A-27
14	1020	99.5	290	1	US-08-485-286-27
15	1020	99.5	290	6	5248606-4
16	1010	98.5	267	1	US-08-218-303-16
17	1010	98.5	267	2	US-08-338-793D-61
18	1010	98.5	267	4	US-09-538-873-1
19	930.5	90.8	540	1	US-08-378-761A-77
20	930.5	90.8	540	1	US-08-485-286-77
21	342	33.4	247	1	US-08-488-113B-6
22	342	33.4	247	1	US-08-477-484B-6
23	342	33.4	247	2	US-08-646-360-6
24	342	33.4	247	3	US-08-839-765-6
25	342	33.4	247	3	US-09-136-389-6
26	342	33.4	247	4	US-09-610-838-6
27	342	33.4	267	1	US-08-378-761A-74

28	342	33.4	267	1	US-08-485-286-74	Sequence 74, Appl
29	342	33.4	289	1	US-07-923-692C-4	Sequence 4, Appl
30	342	33.4	289	1	US-08-184-237-4	Sequence 4, Appl
31	342	33.4	289	2	US-08-482-920-4	Sequence 4, Appl
32	342	33.4	289	3	US-08-484-341-4	Sequence 4, Appl
33	342	33.4	289	3	US-08-483-502-4	Sequence 4, Appl
34	342	33.4	289	4	US-09-726-651A-4	Sequence 4, Appl
35	341.5	33.3	282	1	US-08-324-301-15	Sequence 15, Appl
36	329.5	32.1	250	1	US-08-378-761A-71	Sequence 71, Appl
37	329.5	32.1	250	1	US-08-485-286-71	Sequence 71, Appl
38	323.5	31.6	251	4	US-09-538-873-3	Sequence 3, Appl
39	312.5	30.5	255	1	US-07-901-707-6	Sequence 6, Appl
40	312.5	30.5	255	1	US-07-988-430-6	Sequence 6, Appl
41	312.5	30.5	255	1	US-08-425-336-6	Sequence 6, Appl
42	312.5	30.5	255	5	PCT-US92-09487-6	Sequence 6, Appl
43	312	30.4	248	3	US-08-902-486-7	Sequence 7, Appl
44	312	30.4	290	1	US-08-245-754A-2	Sequence 2, Appl
45	312	30.4	290	2	US-08-597-731-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-356-786-8  
; Sequence 8, Application US/08356786  
; Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
TITLE OF INVENTION: Marker  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS: 16  
ADDRESSES: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,786  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-786-8

Query Match 100.0%; Score 1025; DB 2; Length 268;  
Best Local Similarity 100.0%; Pred. No. 66-112;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFIL 60  
DB 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFIL 60  
QY 61 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAF 120  
DB 61 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAF 120  
QY 121 GGNDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQPLTLARSFIIQIMISEAA 180  
DB 121 GGNDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQPLTLARSFIIQIMISEAA 180  
QY 181 RFOYIEGEMTRIRYNRRS 199  
DB 181 RFOYIEGEMTRIRYNRRS 199

## RESULT 2

US-08-356-786-10  
; Sequence 10, Application US/08356786  
; Patent No. 5877305

## GENERAL INFORMATION:

APPLICANT: Huston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
TITLE OF INVENTION: Marker  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,786  
FILING DATE:

## CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100

## INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:  
LENGTH: 534 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-786-10

Query Match 100.0%; Score 1025; DB 2; Length 534;  
Best Local Similarity 100.0%; Pred. No. 1.7e-111;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFIL 60  
DB 3 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFIL 62  
QY 61 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAF 120

DB 63 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAF 122  
QY 121 GGNDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQPLTLARSFIIQIMISEAA 180  
DB 123 GGNDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQPLTLARSFIIQIMISEAA 182  
QY 181 RFOYIEGEMTRIRYNRRS 199  
DB 183 RFOYIEGEMTRIRYNRRS 201

## RESULT 3

US-07-901-707-1  
; Sequence 1, Application US/07901707  
; Patent No. 5376546

## GENERAL INFORMATION:

APPLICANT: Bernhardt, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Steve F.  
APPLICANT: Lane, Julie A.  
TITLE OF INVENTION: Materials Comprising and Methods of  
TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/901,707  
FILING DATE: 19920619  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5376546and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27129/30910  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-5750  
TELEX: 25-3856

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-901-707-1

Query Match 99.5%; Score 1020; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 2.3e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILV 61  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILV 60  
QY 62 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAF 121  
DB 61 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAF 120  
QY 122 GGNDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQPLTLARSFIIQIMISEAA 181

100

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 11.5641 Seconds  
(without alignments)  
731.761 Million cell updates/sec

Title: US-10-083-336A-10

Perfect score: 1029

Sequence: 1 MIFPKQYPIINFTTAGATVQ.....RFQYIEGEMTRIRYNRSA 200

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

1: /cgn2\_6/prodata/1/iaa/5A COMB.pcp.\*

2: /cgn2\_6/prodata/1/iaa/5B COMB.pcp.\*

3: /cgn2\_6/prodata/1/iaa/6A COMB.pcp.\*

4: /cgn2\_6/prodata/1/iaa/6B COMB.pcp.\*

5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pcp.\*

6: /cgn2\_6/prodata/1/iaa/backfile1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1029	100.0	268	2	US-08-356-786-8
2	1029	100.0	534	2	US-08-356-786-10
3	1024	99.5	267	1	US-07-901-707-1
4	1024	99.5	267	1	US-07-988-430-1
5	1024	99.5	267	1	US-08-425-336-1
6	1024	99.5	267	1	US-08-488-113B-1
7	1024	99.5	267	1	US-08-477-484B-1
8	1024	99.5	267	2	US-08-646-360-1
9	1024	99.5	267	3	US-08-839-765-1
10	1024	99.5	267	3	US-09-136-389-1
11	1024	99.5	267	4	US-09-610-838-1
12	1024	99.5	267	5	PCT-US92-09487-1
13	1024	99.5	230	1	US-08-378-761A-27
14	1024	99.5	230	1	US-08-485-286-27
15	1024	99.5	230	6	5248606-4
16	1014	98.5	267	1	US-08-218-303-16
17	1014	98.5	267	2	US-08-338-793D-61
18	1014	98.5	267	4	US-09-538-873-1
19	934.5	90.8	540	1	US-08-378-761A-77
20	934.5	90.8	540	1	US-08-485-286-77
21	342	33.2	247	1	US-08-488-113B-6
22	342	33.2	247	1	US-08-477-484B-6
23	342	33.2	247	1	US-08-646-360-6
24	342	33.2	247	3	US-08-839-765-6
25	342	33.2	247	3	US-09-136-389-6
26	342	33.2	247	4	US-09-610-838-6
27	342	33.2	267	1	US-08-378-761A-74

28	342	33.2	267	1	US-08-485-286-74
29	342	33.2	289	1	US-07-923-692C-4
30	342	33.2	289	1	US-08-184-237-4
31	342	33.2	289	2	US-08-482-920-4
32	342	33.2	289	3	US-08-484-341-4
33	342	33.2	289	3	US-08-483-502-4
34	342	33.2	289	4	US-09-726-651A-4
35	341.5	33.2	282	1	US-08-324-301-15
36	329.5	32.0	250	1	US-08-378-761A-71
37	329.5	32.0	250	1	US-08-485-286-71
38	323.5	31.4	251	4	US-09-538-873-3
39	312.5	30.4	255	1	US-07-901-707-6
40	312.5	30.4	255	1	US-07-988-430-6
41	312.5	30.4	255	1	US-08-425-336-6
42	312.5	30.4	255	5	PCT-US92-09487-6
43	312	30.3	248	3	US-08-902-486-7
44	312	30.3	290	1	US-08-245-754A-2
45	312	30.3	290	2	US-08-597-731-2

## ALIGNMENTS

RESULT 1  
US-08-356-786-8  
; Sequence B. Application US/08356786  
; Patent No. 5877305

### GENERAL INFORMATION:

APPLICANT: Huston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
TITLE OF INVENTION: Marker  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESS: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356.786  
FILING DATE:

### CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992

### ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053

### TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100

### INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-356-786-8

Query Match 100.0%; Score 1029; DB 2; Length 268;

Best Local Similarity 100.0%; Pred. No. 2.2e-112;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 60  
Db 1 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 60  
QY 61 VELSNAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAF 120  
Db 61 VELSNAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAF 120  
QY 121 GGYDRLEQLAGNLRENIELGNGLPLEAISAALYYSTGGTQLPTLARSFICIQMISEAA 180  
Db 121 GGYDRLEQLAGNLRENIELGNGLPLEAISAALYYSTGGTQLPTLARSFICIQMISEAA 180  
QY 181 RFOYIEGEMTRIRYNRRA 200  
Db 181 RFOYIEGEMTRIRYNRRA 200

## RESULT 2

US-08-356-786-10

; Sequence 10, Application US/08356786

; Patent No. 5677305

; GENERAL INFORMATION:

; APPLICANT: Huston, James S.

; APPLICANT: Oppermann, Hermann

; APPLICANT: Houston, L. L.

; APPLICANT: Ring, David B.

; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer

; TITLE OF INVENTION: Marker

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, &amp; Thibault

; STREET: Exchange Place, 53 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/356,786

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/831,967

; FILING DATE: 06-FEB-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Pitcher, Edmund R.

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: CEP-053

; TELEPHONE: (617) 248-7000

; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 534 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-356-786-10

Query Match

Best Local Similarity 100.0%; Score 1029; DB 2; Length 534;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 60  
Db 3 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 62  
QY 61 VELSNAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAF 120

Db 63 VELSNAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAF 122  
QY 121 GGYDRLEQLAGNLRENIELGNGLPLEAISAALYYSTGGTQLPTLARSFICIQMISEAA 180  
Db 123 GGYDRLEQLAGNLRENIELGNGLPLEAISAALYYSTGGTQLPTLARSFICIQMISEAA 182  
QY 181 RFOYIEGEMTRIRYNRRA 200  
Db 183 RFOYIEGEMTRIRYNRRA 202

## RESULT 3

US-07-901-707-1

; Sequence 1, Application US/07901707

; Patent No. 5376546

; GENERAL INFORMATION:

; APPLICANT: Bernhard, Susan L.

; APPLICANT: Better, Marc D.

; APPLICANT: Carroll, Steve F.

; APPLICANT: Lane, Julie A.

; TITLE OF INVENTION: Materials Comprising and Methods of

; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins

; NUMBER OF SEQUENCES: 57

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marchall, O'Toole, Gerstein, Murray &amp;

; ADDRESSEE: Bicknell, Street

; STREET: Two First National Plaza, 20 South Clark

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60603

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/901,707

; FILING DATE: 19920619

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/787,567

; FILING DATE: 04-NOV-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5376546and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 27129/30910

; TELEPHONE: (312) 346-5750

; TELEFAX: (312) 984-5750

; TELETYPE: 25-3856

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 267 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-07-901-707-1

Query Match

Best Local Similarity 100.0%; Score 1024; DB 1; Length 267;

Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 61  
Db 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 60  
QY 62 VELSNAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAF 121  
Db 61 VELSNAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAF 120  
QY 122 GGYDRLEQLAGNLRENIELGNGLPLEAISAALYYSTGGTQLPTLARSFICIQMISEAA 181

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:35 ; Search time 10.7687 Seconds  
(without alignments)  
1777.145 Million cell updates/sec

Title: US-10-083-336A-5

Perfect score: 1025

Sequence: 1 MIFPKQPIINFITAGATVQ.....ARFQVIGEMTRIRYNRRS 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	99.5	576	1 RLCSA	ricin D precursor
2	930.5	90.8	564	1 RLCSAG	agglutinin precursor
3	353.5	34.5	528	2 S32431	abrin-d precursor
4	345.5	34.5	562	2 S16022	abrin-c precursor
5	345	33.7	527	2 S32430	abrin-b precursor
6	342	33.4	289	1 RLTTZT	RNA N-glycosidase
7	338	33.0	247	2 JUT0393	karasurin - Mongol
8	338	33.0	247	2 JCS032	karasurin-B - Tric
9	336.5	32.8	289	2 JCS606	karasurin C - Tric
10	336.5	32.1	528	1 TLZLSA	abrin (clone 7.2)
11	329.5	30.0	278	2 S23519	abrin-a precursor
12	307.5	29.3	250	2 JN0108	beta-luffin - smoo
13	300.5	29.3	250	2 JN0108	luffin-b - smoo
14	291	28.4	570	2 S26227	agglutinin I precu
15	286	27.9	254	2 P00018	mistletoe lectin I
16	284.5	27.8	277	2 S24494	RNA N-glycosidase
17	279	27.2	286	2 S25560	RNA N-glycosidase
18	278	27.1	245	2 JCL4840	RNA N-glycosidase
19	277	27.0	286	1 RLPU6G	RNA N-glycosidase
20	272	26.5	286	2 JCL4235	RNA N-glycosidase
21	242.5	23.7	316	2 JUT0753	RNA N-glycosidase
22	197.5	19.3	294	2 S28421	RNA N-glycosidase
23	182	17.8	313	2 S17757	RNA N-glycosidase
24	180.5	17.6	261	2 J80401	antiviral protein
25	178	17.4	278	2 A39817	RNA N-glycosidase
26	156	15.2	272	2 JCL4811	betavulgin - beet
27	149.5	14.6	289	2 T12573	RNA N-glycosidase
28	140	13.7	280	1 RL6H	RNA N-glycosidase
29	137	13.4	275	2 S33631	tritin - wheat

## ALIGNMENTS

### RESULT 1

#### RLCSD

ricin D precursor - castor bean

N:Contains: RNA N-glycosidase (EC 3.2.2.22)

C:Species: Ricinus communis (castor bean)

C>Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text change 16-Jul-1999

C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.

Nucleic Acids Res. 13, 8019-8033, 1985

A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A:Reference number: A24041; MUID:86067214; PMID:2999712

A:Accession: A24041

A:Molecule type: DNA

A:Residues: 1-576 <HAL>

A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R:Tregear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin gene

A:Reference number: S20513; MUID:92163016; PMID:1371405

A:Accession: S20513

A:Molecule type: DNA

A:Residues: 1-576 <TR>

A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R:Lamb, F.I.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A:Title: Nucleotide sequence of cloned cDNA coding for preprorcin.

A:Reference number: A24614; MUID:85179479; PMID:3838723

A:Accession: A24614

A:Molecule type: mRNA

A:Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>

A:Cross-references: GB:X02388; NID:g21077; PIDN:CAA26230.1; PID:g21078

R:Yoshitake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A:Title: Isolation and sequences of peptic peptides, and the complete sequence of Ile cha

A:Reference number: A03372

A:Accession: A03372

A:Molecule type: protein

A:Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 272-283, 'L', 285-288, 290-302 <YOS>

A:Note: this paper cites the others in the series providing experimental details for the

R:Araki, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1985

A:Title: Reversed amino acid sequence of the B-chain of ricin D due to loss of tryptophan

A:Reference number: A24010

A:Accession: A24010

A:Molecule type: protein

A:Residues: 315-383, 'PS', 386-576 <ARA>

R:Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A:Title: Primary structure of Ala chain of ricin D.

A:Reference number: A03374

A:Accession: A03374

A:Molecule type: protein

A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405, 'V', 529-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F' <FUN>  
 A:Note: This paper, one of a series, summarizes the experimental details for the determination of the complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.  
 R:Ready, M.P.; Kim, Y.; Robertus, J.D.  
 Proteins 10, 270-278, 1991  
 A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism of action of ricin.  
 A:Reference number: A48237; PMID:91352006; PMID:1881883  
 A:Contents: annotation; active site  
 R:Rutenber, E.; Robertus, J.D.  
 Proteins 10, 260-269, 1991  
 A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.  
 A:Reference number: A48238; PMID:91352005; PMID:1881882  
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms  
 R:Katzin, B.J.; Collins, E.J.; Robertus, J.D.  
 Proteins 10, 251-259, 1991  
 A:Title: Structure of ricin A-chain at 2.5 angstroms.  
 A:Reference number: A48239; PMID:91352004; PMID:1881881  
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms  
 C:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which inactivates the 60S ribosomal subunit.  
 C:Comment: The A chain inhibits protein synthesis; it inactivates the 60S ribosomal subunit.  
 C:Comment: This protein is cytotoxic and very poisonous to animals.  
 C:Superfamily: ricin; rRNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed  
 F:1-35/Domain: signal sequence #status predicted <SIG>  
 F:36-302/Product: ricin D chain A #status experimental <ACH>  
 F:46-293/Domain: rRNA N-glycosidase homology <RNG>  
 F:315-576/Product: ricin D chain B #status experimental <BCH>  
 F:331-373, 374-414, 417-455, 462-497, 501-540, 543-576/Region: 40-residue repeats  
 F:45, 409, 449/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F:115, 158, 243, 244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:212/Active site: Glu #status experimental  
 F:215/Active site: Arg #status predicted  
 F:294-318, 334-353, 377-394, 465-478, 504-521/Disulfide bonds: #status experimental  
 F:336, 349, 360/Binding site: N-acetylglucosamine (Asp, Glu, Asn) #status experimental  
 F:548, 569/Binding site: N-acetylglucosamine (Asp, Asn) #status experimental  
 Query Match 99.5%; Score 1020; DB 1; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 5e-84; Indels 0; Gaps 0;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 61  
 Db 36 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 95  
 QY 62 ELSNHAELSVTLLDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 121  
 Db 96 ELSNHAELSVTLLDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 155  
 QY 122 GNYDRLEQLAGNRLNIELGNGLPGLAEISALYYSTGGTQPLTARSFFICIQMISEAAR 181  
 Db 156 GNYDRLEQLAGNRLNIELGNGLPGLAEISALYYSTGGTQPLTARSFFICIQMISEAAR 215  
 QY 182 FOYIEGEMTRIRYNRRS 199  
 Db 216 FOYIEGEMTRIRYNRRS 233  
 RESULT 2  
 RLCSAG  
 agglutinin precursor - castor bean  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Ricinus communis (castor bean)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C:Accession: A24261; A24210  
 R:Robertus, L.M.; Lamb, I.F.; Pappin, D.J.C.; Lord, J.M.  
 J. Biol. Chem. 260, 15682-15686, 1985  
 A:Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.  
 A:Reference number: A24261; PMID:86059449; PMID:2999130  
 A:Accession: A24261  
 A:Molecule type: mRNA  
 A:Residues: 1-564 <ROB>  
 A:Cross-references: GB:M12089; NID:g169700; PIDN:AAA33869.1; PID:g169701

R:Araki, T.; Yoshioka, Y.; Funatsu, G.  
 Biochim. Biophys. Acta 872, 277-285, 1986  
 A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.  
 A:Reference number: A24210  
 A:Accession: A24210  
 A:Molecule type: Protein  
 A:Residues: 303-325, 'F', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 553-554, 'W', 556, 'H', 557-560, 'LI', 563-564, 'F' <FUN>  
 C:Comment: This protein has strong agglutinating activity and weak cytotoxicity compared with ricin.  
 C:Superfamily: ricin; rRNA N-glycosidase homology  
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-290/Product: agglutinin chain A #status predicted <ACH>  
 F:303-564/Product: agglutinin chain B #status experimental <BCH>  
 F:319-361, 362-402, 405-443, 450-485, 489-528, 531-564/Region: 40-residue repeats  
 F:34, 259/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:104, 147, 231, 232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F:200, 203/Active site: Glu, Arg #status predicted  
 F:282-306, 322-341, 365-382, 453-466, 492-509/Disulfide bonds: #status predicted  
 F:324, 337, 348/Binding site: N-acetylglucosamine (Asp, Glu, Asn) #status predicted  
 F:397, 437/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F:536, 557/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 Query Match 90.8%; Score 930.5; DB 1; Length 564;  
 Best Local Similarity 91.9%; Pred. No. 5.6e-76; Indels 1; Gaps 1;  
 Matches 182; Conservative 7; Mismatches 8; Indels 1; Gaps 1;  
 QY 2 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 61  
 Db 25 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 84  
 QY 62 ELSNHAELSVTLLDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 121  
 Db 85 ELSNHAELSVTLLDVTNAYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 144  
 QY 122 GNYDRLEQLAGNRLNIELGNGLPGLAEISALYYSTGGTQPLTARSFFICIQMISEAAR 181  
 Db 145 GNYDRLEQLAGNRLNIELGNGLPGLAEISALYYSTGGTQPLTARSFFICIQMISEAAR 203  
 QY 182 FOYIEGEMTRIRYNRRS 199  
 Db 204 FOYIEGEMTRIRYNRRS 221  
 RESULT 3  
 S32431  
 abrin-d precursor - Indian licorice (fragment)  
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C:Species: Abrus precatorius (Indian licorice)  
 C:Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 01-Aug-1997  
 C:Accession: S32431; S34408  
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A:Title: Primary structure of three distinct isoabrinins determined by cDNA sequencing. Covalent structure of abrin-d chain A.  
 A:Reference number: S32429; PMID:93132798; PMID:8421313  
 A:Accession: S32431  
 A:Molecule type: mRNA  
 A:Residues: 1-528 <HUN>  
 A:Cross-references: GB:M98346  
 R:Hung, C.; Lee, M.; Lee, T.; Lin, J.  
 submitted to the EMBL Data Library, March 1993  
 A:Reference number: S34408  
 A:Accession: S34408  
 A:Molecule type: mRNA  
 A:Residues: 1-169, 'C', 171-320, 'L', 322-528 <HU2>  
 A:Cross-references: GB:M98346  
 C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating the 60S ribosomal subunit, and a B chain, which is essential for toxicity.  
 C:Superfamily: ricin; rRNA N-glycosidase homology  
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; p  
 F:1-251/Product: abrin-d chain A #status predicted <ACH>  
 F:7-246/Domain: rRNA N-glycosidase homology <RNG>  
 F:261-528/Product: abrin-d chain B #status predicted <BCH>

F;283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats  
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
 F;74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F;164,167/Active site: Glu, Arg #status predicted  
 F;200,253,361,401,402/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F;247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted  
 F;288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F;500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 34.5%; Score 353.5; DB 2; Length 528;  
 Best Local Similarity 45.5%; Pred. No. 5.2e-24;  
 Matches 86; Conservative 24; Mismatches 70; Indels 9; Gaps 4;

QY 6 QYPIINFTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNVRVGLPINQRFILVLSN 65  
 Db 1 QDQVIKFTTEGATSSQSKQFIEALRQLRTGG--LIHDIPVLPDPTTVEERNRYITVLSN 58  
 QY 66 HAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNRVYTFAGGNYD 125  
 Db 59 SERESIEVGIDVTNAYVYVYRAGSQSYFL---RDAPASASTYLPFGTQ-RYSRFDGSGY 114  
 QY 126 RLQLAGNLRENIELNGPLLEAISALYYSTGTGTLPLARSFIICMIOMISEARFOYI 185  
 Db 115 DLERWAHTQREISLGLQALTHAIS---FLRSGASNDDEKARTLIVIQMASEARVYI 171  
 QY 186 EGEWRTIR 194  
 Db 172 SNRVGVSI 180

## RESULT 4

S16022  
 abrin-c precursor - Indian licorice  
 N;Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C;Species: Abrus precatorius (Indian licorice)  
 C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999  
 C;Accession: S16022  
 R;Wood, K.A.; Lord, J.M.; Wawrzynczak, E.J.; Piatak, M.  
 Eur. J. Biochem. 198, 723-732, 1991  
 A;Title: Prochrobrin: genomic cloning, characterisation and the expression of the A-chain  
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
 A;Reference number: S16022; MUID:91266957; PMID:2050149  
 A;Accession: S16022  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-562 <WO>  
 A;Cross-references: EMBL:X55667; NID:gl6084; PIDN:CAA39202.1; PID:gl6085  
 C;Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating  
 The A and B chains are linked by a single disulfide bond, which is essential for toxicity  
 C;Superfamily: ricin; rRNA N-glycosidase homology  
 C;Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglytamic acid  
 F;35-285/Product: abrin-c chain A #status predicted <ACH>  
 F;41-280/Domain: rRNA N-glycosidase homology <RNG>  
 F;295-562/Product: abrin-c chain B #status predicted <BCH>  
 F;317-359,360-400,403-441,448-483,487-526,529-562/Region: 40-residue repeats  
 F;35/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
 F;108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F;198,401/Active site: Glu, Arg #status predicted  
 F;234,287,395,435,436/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F;281-303,320-339,363-380,451-464,490-507/Disulfide bonds: #status predicted  
 F;322,346/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F;534,555/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 34.5%; Score 353.5; DB 2; Length 562;  
 Best Local Similarity 45.5%; Pred. No. 5.6e-24;  
 Matches 86; Conservative 24; Mismatches 70; Indels 9; Gaps 4;

QY 6 QYPIINFTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNVRVGLPINQRFILVLSN 65  
 Db 35 QDQVIKFTTEGATSSQSKQFIEALRQLRTGG--LIHDIPVLPDPTTVEERNRYITVLSN 92  
 QY 66 HAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNRVYTFAGGNYD 125  
 Db 93 SERESIEVGIDVTNAYVYVYRAGSQSYFL---RDAPASASTYLPFGTQ-RYSRFDGSGY 148

QY 126 RLQLAGNLRENIELNGPLLEAISALYYSTGTGTLPLARSFIICMIOMISEARFOYI 185  
 Db 149 DLERWAHTQREISLGLQALTHAIS---FLRSGASNDDEKARTLIVIQMASEARVYI 205  
 QY 186 EGEWRTIR 194  
 Db 206 SNRVGVSI 214

## RESULT 5

S32430  
 abrin-b precursor - Indian licorice (fragment)  
 N;Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C;Species: Abrus precatorius (Indian licorice)  
 C;Date: 30-Sep-1993 #sequence\_revision 01-Aug-1997 #text\_change 20-Aug-1999  
 C;Accession: S32430; JCI399  
 R;Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A;Title: Primary structure of three distinct isoabrinins determined by cDNA sequencing. Cor  
 A;Reference number: S32429; MUID:93132798; PMID:8421313  
 A;Accession: S32430  
 A;Molecule type: mRNA  
 A;Residues: 1-527 <HUN>  
 A;Cross-references: GB:M98345; NID:gl66296; PIDN:AAA32625.1; PID:gl66297  
 R;Kimura, M.; Sumizawa, T.; Funatsu, G.  
 Biosci. Biotechnol. Biochem. 57, 166-169, 1993  
 A;Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic  
 A;Reference number: JCI398; MUID:93169023; PMID:7763422  
 A;Accession: JCI399  
 A;Molecule type: protein  
 A;Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'M', 427, 'D', 429-430,  
 A;Experimental source: seed  
 C;Superfamily: ricin; rRNA N-glycosidase homology  
 C;Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; py  
 F;1-250/Product: abrin-b chain A #status predicted <ACH>  
 F;7-245/Domain: rRNA N-glycosidase homology <RNG>  
 F;260-527/Product: abrin-b chain B #status experimental <BCH>  
 F;282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats  
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted  
 F;74,113,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F;110,360,400/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F;163,166/Active site: Glu, Arg #status predicted  
 F;246-268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted  
 F;287,311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
 F;499,520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 33.7%; Score 345; DB 2; Length 527;  
 Best Local Similarity 45.5%; Pred. No. 3e-23;  
 Matches 87; Conservative 22; Mismatches 72; Indels 10; Gaps 4;

QY 6 QYPIINFTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNVRVGLPINQRFILVLSN 65  
 Db 1 QDQVIKFTTEGATSSQSKQFIEALRQLRTGG--LIHDIPVLPDPTTVEERNRYITVLSN 58  
 QY 66 HAELSVTLALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNRVYTFAGGNYD 125  
 Db 59 SDTESIEAGIDVSNAYVYVYRAGNSYFL---RDAPTSASRYLFTGTQ-QYSLRFNGSYI 114  
 QY 126 RLQLAGNLRENIELNGPLLEAISALYYSTGTGTLPLARSFIICMIOMISEARFOYI 185  
 Db 115 DLERWAHTQREISLGLQALTHAISFL---QSGTDDQEIARTLIVIQMASEARVYI 170  
 QY 186 EGEWRTIRYN 196  
 Db 171 SYRVGVSI 181

## RESULT 6

RLT2T  
 rRNA N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolian sr  
 N;Alternate names: alpha-TCS; type I ribosome-inactivating protein  
 C;Species: Trichosanthes kirilowii (Mongolian snake-gourd)



C;Date: 30-Sep-1988 #sequence\_revision 26-Jan-1996 #text\_change 23-Mar-2001  
C;Accession: JT0566; A36274; JCI093; A36273; JT0003  
R;Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.  
Gene 97, 267-272, 1991  
A;Title: Cloning of trichosanthin cDNA and its expression in *Escherichia coli*.  
A;Reference number: JT0566; MUID:91153657; PMID:1999291  
A;Accession: JT0566  
A;Molecule type: mRNA  
A;Residues: 1-289 <SHA>  
A;Cross-references: GB:M34858; NID:gi705336; PIDN:AAA34207.1; PID:gi70537  
A;Experimental source: tuber  
R;Chow, T.P.; Feldman, R.A.; Lovett, M.; Piatak, M.  
J. Biol. Chem. 265, 8670-8674, 1990  
A;Title: Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I rib  
A;Reference number: A36274; MUID:90256790; PMID:2341400  
A;Accession: A36274  
A;Molecule type: DNA  
A;Residues: 1-233, 'T', 235-246, 'M', 248-289 <CHO>  
A;Cross-references: GB:J05434; NID:gi705334; PIDN:AAA34206.1; PID:gi70535  
R;Zheng, H.G.; Wang, B.; Shao, P.Z.; Yang, X.R.  
Acta Genet. Sin. 21, 42-51, 1994  
A;Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.  
A;Reference number: JCI093; MUID:94271613; PMID:8003348  
A;Accession: JCI093  
A;Molecule type: DNA  
A;Residues: 1-72, 'V', 74-90, 'S', 92-233, 'T', 235-267, 'D', 269-289 <ZHE>  
A;Cross-references: GB:S70176; NID:G547148; PIDN:AAB31048.1; PID:G547149  
R;Collins, E.J.; Robertus, J.D.; LoPresti, M.; Stone, K.L.; Williams, K.R.; Wu, P.; Hwan  
J. Biol. Chem. 265, 8665-8669, 1990  
A;Title: Primary amino acid sequence of alpha-trichosanthin and molecular models for ab  
A;Reference number: A36273; MUID:90256789; PMID:2341399  
A;Accession: A36273  
A;Molecule type: protein  
A;Residues: 24-270 <CO>  
R;Wang, Y.; Qian, R.Q.; Gu, Z.W.; Jin, S.W.; Zhang, L.Q.; Xia, Z.X.; Tian, G.Y.; Ni, C.Z  
Pure Appl. Chem. 58, 789-798, 1986  
A;Title: Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application  
A;Reference number: JT0003  
A;Accession: JT0003  
A;Molecule type: protein  
A;Residues: 24-56, 'L', 58-59, 'I', 61-71, 'I', 73-81, 85-86, 'L', 88-92, 'DAGLPNNAVL', 93-142, 'GL'  
A;Experimental source: tuber  
R;Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, July 1994  
A;Reference number: A67091; PDB:1MRJ  
A;Contents: annotation; X-ray crystallography, 1.6 angstroms, with adenine, residues 24-  
R;Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, July 1994  
A;Reference number: A67092; PDB:1MRK  
A;Contents: annotation; X-ray crystallography, 1.6 angstroms, with formycin, residues 24  
R;Xiong, J.P.; Xia, Z.X.; Wang, Y.  
submitted to the Brookhaven Protein Data Bank, December 1994  
A;Reference number: A66711; PDB:1RCS  
A;Contents: annotation; X-ray crystallography, 1.7 angstroms, with NADPH, residues 24-27  
R;Xiong, J.P.; Xia, Z.X.; Wang, Y.  
Nat. Struct. Biol. 1, 695-700, 1994  
A;Title: Crystal structure of trichosanthin-NADPH complex at 1.7 Angstroms resolution re  
A;Reference number: A58622; MUID:95360714; PMID:7634073  
A;Contents: annotation; X-ray crystallography, 1.7 angstroms  
C;Comment: Alpha-trichosanthin has been used to induce abortions.  
C;Genetics:  
A;Gene: tcs  
C;Function:  
A;Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA then  
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C;Keywords: abortifacient; glycosidase; hydrolase; root; toxin  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-270/Product: trichosanthin alpha #status experimental <MAT>  
F;271-289/Domain: rRNA N-glycosidase homology <RNG>  
F;93,183,186/Active site: Tyr, Glu, Arg #status predicted <CTP>  
Query Match 33.4%; Score 342; DB 1; Length 289;

Best Local Similarity 39.5%; Pred. No. 2.6e-23;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
QY 10 INFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFILLVLSNHAEL 69  
DB 25 VSFRLSGATSSYGVFISNLRKALPYERKL-YDIPLL--RSLPGSQRYALHILNTYADE 81  
QY 70 SVTLALDVTNAYVVGVRAGNSAYFFPHDPNQEDA-BAITHLFTDVQNRVTFAPFGNVDRL 128  
DB 82 TTSVAIDVTNVYVVGVRAGDTSYFF---NEASATEAAKYVFKDAMRKVTLTPYSGNYERLQ 138  
QY 129 QLAGNLRNIEIENGFLPEALISALYYSTGGTQTLPLARSFLICQMISEAARFQYIEGE 188  
DB 139 TAAGKIRENIPLGLPALDSAITTLFYNNAN-----SAASALMVLIIQSTSEAAARYKFIEQQ 193  
QY 189 MRTRI 193  
DB 194 ICKRV 198  
RESULT 7  
JU0393  
karasurin - Mongolian snake-gourd  
C;Species: Trichosanthes kirilowii (Mongolian snake-gourd)  
C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 12-Apr-1995  
C;Accession: JU0393; PS0163  
R;Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogihara, Y.  
Chem. Pharm. Bull. 39, 1244-1249, 1991  
A;Title: The complete amino acid sequence of an abortifacient protein, karasurin.  
A;Reference number: JU0393; MUID:92005921; PMID:1914000  
A;Accession: JU0393  
A;Molecule type: protein  
A;Residues: 1-247 <TOY>  
A;Note: a sequence which lacks Ala-247 is also shown in this publication  
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
C;Keywords: abortifacient  
F;1-243/Domain: rRNA N-glycosidase homology <RNG>  
Query Match 33.0%; Score 338; DB 2; Length 247;  
Best Local Similarity 40.0%; Pred. No. 4.9e-23;  
Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5;  
QY 10 INFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINORFILLVLSNHAEL 69  
DB 2 VSFRLSGATSSYGVFISNLRKALPYERKL-YDIPLL--RSLPGSQRYALHILNTYADE 58  
QY 70 SVTLALDVTNAYVVGVRAGNSAYFFPHDPNQEDA-BAITHLFTDVQNRVTFAPFGNVDRL 128  
DB 59 TTSVAIDVTNVYVVGVRAGDTSYFF---NEASATEAAKYVFKDAMRKVTLTPYSGNYERLQ 115  
QY 129 QLAGNLRNIEIENGFLPEALISALYYSTGGTQTLPLARSFLICQMISEAARFQYIEGE 188  
DB 116 IAAGKIRENIPLGLPALDSAITTLFYNNAN-----SAASALMVLIIQSTSEAAARYKFIEQQ 170  
QY 189 MRTRI 193  
DB 171 ICKRV 175  
RESULT 8  
JCS032  
karasurin-B - Trichosanthes kirilowii var. japonica  
C;Species: Trichosanthes kirilowii var. japonica  
C;Date: 27-Feb-1997 #sequence\_revision 27-Feb-1997 #text\_change 23-May-1997  
C;Accession: JCS032  
R;Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.  
Biol. Pharm. Bull. 19, 1485-1489, 1996  
A;Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka  
A;Reference number: JCS032; MUID:97108848; PMID:8951169  
A;Accession: JCS032  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-247 <KON>

C;Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyze the N-glycosidic bond cleavage of 28S rRNA. C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology F;4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match	33.0%;	Score 338;	DB 2;	Length 247;
Best Local Similarity	40.0%;	Pred. No. 4.9e-23;		
Matches	74;	Conservative 46;	Mismatches 53;	Indels 12; Gaps 5;
QY	10	INFTAGATQSQYNTFRVAVRGRLTTTGADVDRHEIPLVLPNRVGLPINOFRFIILVELSNHAEL	69	
Db	2	VSRFLSGATSSSCYGVFSNRKALPYEKL-YDIPLL--RSTLPQSQYALHINNYADE	58	
QY	70	SVTLALDVNTNAYVVGYPAGNSAYFFHPDNOBDA-EAITHLETDVQNRVYTFAGGNYDRLE	128	
Db	59	TISVAIDVTNNVYMGYPAGDTSYFF--NEASATEAAKYVEFKARVKVTLPSGNYERLQ	115	
QY	129	QLAGNLRNTELGNGPLEEATISALYYVSTGQTLPRLARSIICTIOMISEARFQYIEGE	188	
Db	116	IAGKIRENTEPLGLPALDSAITLTFYNN-----SASALMVLITQSTSEARFYKFIQQ	170	
QY	189	MRTRI	193	
Db	171	IGKRV	175	

RESULT 9  
JC5606  
karasurin C - *Trichosanthes kirilowii* var. *japonica*  
N; Contains: Karasurin A  
C; Species: *Trichosanthes kirilowii* var. *japonica*  
C; Date: 23-Sep-1997 #sequence\_revision 23-Sep-1997 #text\_change 19-Jul-2002  
C; Accession: JC5606; JC5033  
R; Mizukami, H.; Iida, K.; Kondo, T.; Ogiwara, Y.  
Biol. Pharm. Bull. 20, 711-713, 1997  
A; Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating protein  
A; Reference number: JC5606; MUID:97356562; PMID:9212998  
A; Accession: JC5606  
A; Molecule type: DNA  
A; Residues: 1-289 <MIZ>  
A; Cross-residues: DDBJ:AB000666; NID:92329830; PIDN:BAA21786.1; PID:92329831  
R; Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.  
Biol. Pharm. Bull. 19, 1485-1489, 1996  
A; Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and karasurin-C  
A; Reference number: JC5032; MUID:97108848; PMID:8951169  
A; Accession: JC5033  
A; Status: preliminary  
A; Molecule type: protein  
A; Residues: 22-270 <KON>  
C; Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, abortifacient and antiproliferative activities.  
C; Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F; 22-270/Product: karasurin C #status predicted <MAC>  
F; 24-270/Product: karasurin A #status predicted <MAA>  
F; 27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match	33.0%;	Score 338;	DB 2;	Length 289;
Best Local Similarity	40.0%;	Pred. No. 5.9e-23;		
Matches	74;	Conservative 46;	Mismatches 53;	Indels 12; Gaps 5;
QY	10	INFTAGATVQSNTFNIRAVRGRLTTGADVREHIEVLFPNVRVGLPINQRFLVLSNHAEL	69	
Db	25	VSRFLSGATSSSYGVFISNLRKALPYERKL--YDIPLL--RSTLPGSQRYALHILTNVADE	81	
QY	70	SVTLALDVTNAAVVGYSRAGNSAYFFHPDNOEDA-EAITHLFTDVQNVRYTFAGGNYDELE	128	
Db	82	TISVAIDVTNVVVGYSRAGDTSYFF---NEASATEAAKYVPDAKKVTLTPSGNYERLQ	138	
QY	129	QLAGNIRENTELGMPLEBAISALYYSTGTQTPTLARSTICIMISEARFQYIEGE	188	
Db	139	IAAGKIRENIPGLPALDSAITTLTFYNNAN----SAASALMVLTIQSTSEAAKYFTEQQ	193	
QY	189	MRTRI	193	
		:	:	

Db 194 IGKRV 198

## RESULT 10

C39761  
 abrin (clone 7.2) precursor - Indian licorice (fragment)  
 N;Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C;Species: Abrus precatorius (Indian licorice)  
 C;Date: 21-Feb-1992 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
 C;Accession: C39761, S14471  
 R;Evensen, G.; Mathiesen, A.; Sundan, A.  
 J. Biol. Chem. 266, 6848-6852, 1991  
 A;Title: Direct molecular cloning and expression of two distinct abrin A-chains  
 A;Reference number: A39761; PMID:91201329; PMID:2016300  
 A;Accession: C39761  
 A;Molecule type: DNA  
 A;Residues: 1-251 <EVE>  
 R;Evensen, G.; Mathiesen, A.; Sundan, A.  
 submitted to the EMBL Data Library, October 1990  
 A;Description: Direct molecular cloning of two distinct abrin A-chains.  
 A;Reference number: S14471  
 A;Accession: S14471  
 A;Molecule type: DNA  
 A;Residues: 'M', 1-251 <EV2>

A;Residues: 'M', 1-251 &lt;EV2&gt;

A; Cross-references: EMBL:X54872; NID:G16089; PIDN:CAA38654.1.; PID:G16089  
C; Superfamily: ricin; rRNA N-glycosidase homology  
C; Keywords: duplication; glycosidase; hydrolase; lectin; toxin  
F; 1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>  
F; 7-246/Domain: rRNA N-glycosidase homology <RNG>  
F; 74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F; 164,167/Active site: Glu, Arg #status predicted

Query Match	32.8%;	Score 336.5;	DB 2;	Length 251;
Best Local Similarity	44.9%;	Pred. NO. 6.8e-23;		
Matches 83;	Conservative 24;	Mismatches 69;	Indels 9;	Gaps 4;

QY	10	INFTAGATVQSYTNFIRAVRGRLTIGADVRHEIPVLPNVRGLPTNQRPILVETLSNHAEL	69
Db	5	IKFTEGATQSYKGFIEALRELRGG--LIHDIPLVRDPTTVERNRNYITVELNSERE	62
QY	70	SVTLALDVTNAVYVYGRAGNSAYFFHPNQDEAEATHLFTDVQNYRTPAFGNTYRLQ	129
Db	63	SIEVGHDVTNAVYVYRAGSQSYFL---RDAPASASTYLTGTQ-RYSERFPGSYGDRL	118
QY	130	LAGNLFRELTNGPLPEEAISALYVYSTGGTQPLTARSFICICOMISAAAFQIETEGM	189
Db	119	WAHOTREQLSLGLOALTHAIS---FLRSGASNDDEKARTLIIIVOWASAAARYIYISNR	175

Ov 190 RTRTR 194

176 190

## RESULT 11

TZLSA  
 abrin-a precursor - Indian licorice (fragment)  
 N;Contains: rRNA N-glycosidase (EC 3.2.2.22)  
 C;Species: Abrus precatorius (Indian licorice)  
 C;Date: 31-Dec-1993 #sequence revision 01-Aug-1997 #text change 16-Jul-1999  
 C;Accession: S32429; JTQ202; A39761; JCI398; S14472; S24133; S74110; S74111  
 R;Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.  
 J. Mol. Biol. 229, 263-267, 1993  
 A;Title: Primary structure of three distinct isoabrisins determined by cDNA sequencing. Cor  
 A;Reference number: S32429; MUID:93132798; PMID:8421313  
 A;Accession: S32429  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 'E', 2-528 <HUN>  
 A;Cross-references: GB:M98344; NID:g156294; PID:AAA32624.1; PID:g166295  
 A;Note: the coding region for the sequence shown is preceded by an ATG codon  
 A;Note: residues 1-8 were derived from the synthesized primer  
 R;Funatsu, G.; Taguchi, Y.; Kanenosono, M.; Yanaka, M.  
 Agric. Biol. Chem. 52, 1095-1097, 1988

```

10 INFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVELSNHAEL 69
5 IKFTEGATSSQYKQFTEALRERUGG--LHLDIPVLDPDTLQERNRIIIVELSNSTE 62
70 SVTIALDVNTAVYVGRAGNSAYFFH--PDNQEDAEAIHLFTDVQNRYYTFAFGNYDRL 127
63 SIEVGIDVTNAVYVAYRAGTQSYFLRDAPSSASD-----YLFTGT-DQHSLPFYTYGD 116
128 BOLAGNREINTELGNGPLBEAISALYYSTGTQTPILARSFIICQMISEAARFOYIEG 187
117 ERWAHQSRQIPLGLIQALTHGIS--PFRSGNDNEEKARTLIIQWVAEAAFRYISN 173
188 EMRTIR 194
174 RRVSIQ 180

RESULT 12
S23519
beta-luffin - smooth loofah
C:Species: Luffa cylindrica (smooth loofah)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Aug-1999
C:Accession: S23519, S23113
R:Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koiwai, A.
Plant Mol. Biol. 19, 887-889, 1992
A:Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-inactivating
A:Reference number: S23519; MUID:92353400; PMID:1643290
A:Accession: S23519
A:Molecule type: mRNA
A:Residues: 1-278 <KAT>
A:Cross-references: EMBL:X62372; NID:g19149; PIDN:CAA44230.1; PID:g19150
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F:26-264/Domain: rRNA N-glycosidase homology <RNG>

Query Match 30.0%; Score 307.5; DB 2; Length 278;
Best Local Similarity 35.8%; Pred. No. 3.2e-20;
Matches 67; Conservative 45; Mismatches 64; Indels 11; Gaps 3;

QY 10 INFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVELSNHAEL 69
DB 24 VESFLSGADSKSYKFTALRKALPSKEKVSNIPLLPASGA---SRVILMQLSNYDAK 80
QY 70 SVTIALDVNTAVYVGRAGNSAYFFHDPNQEDAEAIHLFTDVQNRYYTFAFGNYDRL 129
DB 81 AITMAIDVTNVYIMGLVNSTSYFF---NESDAKLASQVYFKGSTIVTLPYSGNYERLQN 137
QY 130 LAGNREINTELGNGPLEEAISSALYYVSTGGTQTPILARSFIICQMISEAARFOYIEGEM 189
DB 138 AGKVRKEIPLGFRAFDSAITSLPHYDS-----TAAAGAFVLIIQTAAERFKYIEGOI 192
QY 190 RTRIRYN 196
DB 193 IERIPKN 199

RESULT 13
JN0108
luffin-b - smooth loofah
C:Species: Luffa cylindrica (smooth loofah)
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 07-May-1999
C:Accession: JN0108
R:Islam, M.R.; Hirayama, H.; Funatsu, G.
Agric. Biol. Chem. 55, 229-238, 1991
A:Title: Complete amino acid sequence of luffin-b, a ribosome-inactivating protein from
A:Reference number: JN0108; MUID:91248488; PMID:1368666
A:Accession: JN0108
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-250 <ISL>
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F:5-246/Domain: rRNA N-glycosidase homology <RNG>

Query Match 29.3%; Score 300.5; DB 2; Length 250;

```

```

A:Accession: PD0018
A:Molecule type: protein
A:Residues: 1-254 <ESC>
C:Superfamily: ricin; rRNA N-glycosidase homology
F;7-246/Domain: rRNA N-glycosidase homology <RNG>

Query Match      27.9%; Score 286; DB 2; Length 254;
Best Local Similarity 38.4%; Pred. No. 2.4e-18;
Matches 81; Conservative 36; Mismatches 56; Indels 38; Gaps 10;

QY 14 TAGATVQSYTNIIRAVRGRLITGADVRHRIPIVLPNVRVGLPIN--QRFILVELSNHAELSV 71
Db 9  THQTTCGEYFRITILRDYVSSGS-FSNIPILL-ROSTIPVSDAQRFLVELTNQGDSV 66

QY 72 TLALDVNTAYVYVGYRAGNSAYFFHPDNQDEADAEITHLFTDVQNRYRTFAFGNTDRLLEQLA 131
Db 67 TAAIDVTNAYVYVAYQAGDQSYELR-DAPRGAE--THLFTGT-TRSSLPFNGSYPDLELYA 122

QY 132 GNLRENIELGNGLPEAEISALYYSTGGTQLPTLARSFLICQMISEAARF-----Q 183
Db 123 GH-RDQIPLGIDQLTQSVTLARF---PGSGSTRQARSILILIQMISEAARFNPILWRYRQ 178

QY 184 YIE-----GEMRTIRYN 196
Db 179 YINGSGASFLPDVYMLETSGWQQOSTQVQHS 209

Search completed: February 10, 2004, 16:28:01
Job time : 11.7687 secs

```

Search completed: February 10, 2004, 16:28:01  
Job time : 11.7687 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:55 ; Search time 6.63825 Seconds  
(without alignments)  
1409.756 Million cell updates/sec

Title: US-10-083-336A-5

Perfect score: 1025

Sequence: 1 MIFPKQVPIINFETAGATVQ.....ARFQVIEGEMTRINRRS 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	99.5	576	1 RICI RICCO	P02879 ricinus com
2	930.5	90.8	564	1 AGEL RICCO	P06750 ricinus com
3	353.5	34.5	562	1 ABRG ABRPR	P28590 abrus preca
4	345	33.7	527	1 ABRB ABRPR	Q06077 abrus preca
5	342	33.4	289	1 RIPT TRIKI	P09989 trichosanthe
6	341.5	33.3	282	1 RIPT TRIKI	P98184 bryonia dio
7	338	33.0	289	1 RIPS TRIKI	P24478 trichosanthe
8	330.5	32.2	563	1 NIGB SAMNI	P33183 sambucus ni
9	329.5	32.1	528	1 ABRG ABRPR	P11140 abrus preca
10	312	30.4	290	1 RIPI BRYDI	P31185 bryonia dio
11	300.5	29.3	250	1 RIPI BRYDI	P22851 luffa cylin
12	296.5	28.9	286	1 RIPI LUCFY	Q9194 cucumis fig
13	286	27.9	254	1 MIA VISAL	P81446 viscum albu
14	284.5	27.8	277	1 RIPI LUCFY	Q00465 luffa cylin
15	279	27.2	286	1 RIPI MOMBA	P29339 momordica b
16	277	27.0	286	1 RIPI MOMBA	P16094 momordica c
17	260	25.4	294	1 RIPI TRIAN	P56626 trichosanthe
18	242.5	23.7	316	1 RIPI GELMU	P33186 gelonium mu
19	197.5	19.3	294	1 RIPI PHYAM	Q03464 phytoacca
20	182	17.8	313	1 RIPI PHYAM	P10297 phytoacca
21	180.5	17.6	261	1 RIPS PHYAM	P23339 phytoacca
22	175	17.4	278	1 RIPI MIRJA	P21326 mirabilis j
23	140	13.7	280	1 RIPI HORVU	P04399 hordeum vul
24	133	13.0	280	1 RIPI HORVU	P22444 hordeum vul
25	131	12.8	253	1 RIPI SAPOF	Q41391 saponaria o
26	127	12.4	253	1 RIPI SAPOF	Q41389 saponaria o
27	126	12.3	310	1 RIPI PHYAM	Q40772 phytoacca
28	124	12.1	299	1 RIPI SAPOF	P20656 saponaria o
29	123.5	12.0	293	1 RIPI DIACA	P24476 dianthus ca
30	121	11.8	292	1 RIPI SAPOF	P27559 saponaria o
31	111	10.8	319	1 SLTA BP933	P09385 bacterioph
32	110	10.7	236	1 RIPI SAPOF	P27560 saponaria o
33	109.5	10.7	300	1 RIPI MAIZE	P25891 zea mays (m

34	107.5	10.5	301	1 RIPX MAIZE	P28522 zea mays (m
35	103	10.0	304	1 RIP9 MAIZE	P25892 zea mays (m
36	102	10.0	560	1 JI60 HORVU	Q00531 hordeum vul
37	91	8.9	315	1 SLTA BP919	P08026 bacterioph
38	91	8.9	315	1 SLTA BP930	P10149 bacterioph
39	85.5	8.3	1165	1 SVV AQUAE	O67411 aquifex aeo
40	85	8.3	609	1 FISH TREPA	O83746 treponema p
41	82.5	8.0	492	1 AERA AERTR	P09166 aeromonas t
42	81	7.9	305	1 Y348 MYCPN	P75205 mycoplasma
43	78.5	7.7	485	1 AERS AERHY	Q06306 aeromonas h
44	78.5	7.7	493	1 AERA AERHY	P09167 aeromonas h
45	76	7.4	157	1 RIP4 SAPOF	P27561 saponaria o

## ALIGNMENTS

RESULT 1  
ID RICI RICCO STANDARD; PRT; 576 AA.  
AC P02879; P02880;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase)  
DE (EC 3.2.2.22); Ricin B chain].  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustersids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86067214; PubMed=2999712;  
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,  
RA Weaver R.F.;  
RT "Genomic cloning and characterization of a ricin gene from Ricinus  
communis";  
RL Nucleic Acids Res. 13:8019-8033(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92163016; PubMed=1371405;  
RA Trisegar J.W., Roberts L.M.;  
RT "The lectin gene family of Ricinus communis: cloning of a functional  
RT ricin gene and three lectin pseudogenes.";  
RL Plant Mol. Biol. 18:515-525(1992).  
RN [3]  
RP SEQUENCE OF 12-576 FROM N.A.  
RX MEDLINE=85179479; PubMed=3838723;  
RA Lamb A., Roberts L.M., Lord J.M.;  
RT "Nucleotide sequence of cloned cDNA coding for preproricin.";  
RL Eur. J. Biochem. 148:265-270(1985).  
RN [4]  
RP SEQUENCE OF 36-302.  
RX Yoshitake S., Funatsu G., Funatsu M.;  
RT "Isolation and sequences of peptic peptides, and the complete  
RT sequence of ile chain of ricin-D.";  
RL Agric. Biol. Chem. 42:1267-1274(1978).  
RN [5]  
RP SEQUENCE OF 315-576.  
RX Funatsu G., Kimura M., Funatsu M.;  
RT "Primary structure of Ala chain of ricin D.";  
RL Agric. Biol. Chem. 43:2221-2224(1979).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.  
RX MEDLINE=90344223; PubMed=1368517;  
RA Kimura Y., Kusuoku H., Tada M., Takagi S., Funatsu G.;  
RT "Structural analyses of sugar chains from ricin A-chain variant.";  
RL Agric. Biol. Chem. 54:157-162(1990).  
RN [7]  
RP REVIEW.  
RX MEDLINE=21480122; PubMed=11595634;  
RA Olmes S., Kozlov J.V.;

RT "Ricin.";  
 RL Toxicol 39:1723-1728 (2001).  
 RP [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=87165983; PubMed=3558397;  
 RA Monfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B.,  
 RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;  
 RT "The three-dimensional structure of ricin at 2.8 Å.";  
 RL J. Biol. Chem. 262:5398-5403 (1987).  
 RP [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=91352004; PubMed=1881881;  
 RA Katzin B.J., Collins E.J., Robertus J.D.;  
 RT "Structure of ricin A-chain at 2.5 Å.";  
 RL Proteins 10:251-259 (1991).  
 RP [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.  
 RX MEDLINE=91352005; PubMed=1881882;  
 RA Rutenber E., Robertus J.D.;  
 RT "Structure of ricin B-chain at 2.5-Å resolution.";  
 RL Proteins 10:260-269 (1991).  
 RP [11]  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=95082010; PubMed=7990130;  
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,  
 RA Paupit R.A.;  
 RT "X-ray structure of recombinant ricin A-chain at 1.8-Å resolution.";  
 RL J. Mol. Biol. 244:410-422 (1994).  
 RP [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.  
 RX MEDLINE=96374222; PubMed=8780513;  
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,  
 RA Molina-Svinth M.C., Robertus J.D.;  
 RT "Structure and activity of an active site substitution of ricin A chain.";  
 RL Biochemistry 35:11098-11103 (1996).  
 RP [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.  
 RX MEDLINE=97240820; PubMed=9086280;  
 RA Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Milne G.W.,  
 RA Robertus J.D.;  
 RT "Structure-based identification of a ricin inhibitor.";  
 RL J. Mol. Biol. 266:1043-1049 (1997).  
 RP [14]  
 RP MUTAGENESIS.  
 RX MEDLINE=93165632; PubMed=1287657;  
 RA Kin Y., Robertus J.D.;  
 RT "Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography.";  
 RL Protein Eng. 5:775-779 (1992).  
 CC -!- FUNCTION: Ricin is highly toxic to animal cells and to a less extent to plant cells. The A chain is responsible for inhibiting protein synthesis through the catalytic inactivation of 60S ribosomal subunits. It acts as a glycosylase that removes a specific adenine residue from an exposed loop of 28S ribosomal RNA. As this loop is involved in the binding of elongation factors, the modified ribosomes are unable to support protein synthesis. The A chain can inactivate a few thousand ribosomes per minute, thus inactivating them faster than the cell can make new ones. A single A-chain molecule can therefore kill an animal cell. The B chain binds to cell receptors and facilitates the entry into the cell of the A chain; B chains are also responsible for cell agglutination (lectin activity). It binds to beta-D-galactopyranoside moieties.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -!- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 CC -!- DOMAIN: The B chain is composed of two domains, each domain consists of 3 homologous subdomains (alpha, beta, gamma).  
 CC -!- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -!- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).  
 CC -!- DATABASE: NAME=Protein Spotlight;  
 CC NOTE=Issue 31 of February 2003;  
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt031.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; X03179; CAA26939.1; -;  
 DR EMBL; X52908; CAA37095.1; -;  
 DR EMBL; X02388; CAA26230.1; -;  
 DR EMBL; A12892; CAA01058.1; -;  
 DR FIR; A24041; RLCSD.  
 DR PDB; 2AAI; 31-JAN-94.  
 DR PDB; 1AFG; 31-JAN-94.  
 DR PDB; 1FMP; 31-OCT-93.  
 DR PDB; 1IFS; 14-JAN-98.  
 DR PDB; 1IFT; 14-JAN-98.  
 DR PDB; 1IFU; 14-JAN-98.  
 DR PDB; 1FTC; 31-OCT-93.  
 DR PDB; 1OBS; 16-JUN-97.  
 DR PDB; 1OBT; 16-JUN-97.  
 DR PDB; 1BR5; 02-SEP-98.  
 DR PDB; 1BR6; 02-SEP-98.  
 DR PDB; 1IL3; 16-JAN-02.  
 DR PDB; 1IL4; 16-JAN-02.  
 DR PDB; 1IL9; 16-JAN-02.  
 DR GlycoSuiteDB; P02879; -;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS02331; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal; 3D-structure.  
 FT SIGNAL 1 35  
 FT CHAIN 36 302 RICIN A CHAIN.  
 FT PEPTIDE 303 314 LINKER PEPTIDE.  
 FT CHAIN 315 576 RICIN B CHAIN.  
 FT DOMAIN 321 448 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 451 575 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 331 373 1-ALPHA.  
 FT REPEAT 374 414 1-BETA.  
 FT REPEAT 417 449 1-GAMMA.  
 FT REPEAT 462 497 2-ALPHA.  
 FT REPEAT 501 540 2-BETA.  
 FT REPEAT 543 570 2-GAMMA.  
 FT ACT\_SITE 212 212  
 FT DISULFID 294 318 INTERCHAIN.  
 FT DISULFID 334 353  
 FT DISULFID 377 394  
 FT DISULFID 465 478  
 FT DISULFID 504 521  
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 271 271 /FTIG-CAR 000080.  
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (IN MINOR FORM).  
 FT CARBOHYD 449 449 /FTIG-CAR 000081.  
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .).  
 FT CONFLICT 76 76 E -> D (IN REF. 3).  
 FT CONFLICT 551 551 A -> R (IN REF. 3).  
 FT STRAND 43 47  
 FT TURN 49 50

Query Match 99.5%; Score 1020; DB 1; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-87;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILV 61  
 DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILV 95

QY 62 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 121  
 DB 96 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 155

QY 122 GNYDRLEQLAGNLRNLELNGPLNEEALISALYYSTGCTQPLTARSFIICQMISEAAR 181  
 DB 156 GNYDRLEQLAGNLRNLELNGPLNEEALISALYYSTGCTQPLTARSFIICQMISEAAR 215

QY 182 FOYIEGEMRTRIRNRRS 199  
 DB 216 FOYIEGEMRTRIRNRRS 233

RESULT 2  
 AGGL RICCO STANDARD; PRT; 564 AA.

AC P06750;  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].  
 DE Ricinus communis (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 OX NCBI\_TaxID=3988;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86059449; PubMed=2999130;  
 RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;  
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";  
 RL J. Biol. Chem. 260:15682-15686(1985).  
 RN [2]  
 RP SEQUENCE OF 303-564.  
 RC TISSUE=Seed;  
 RA Araki T., Yoshioka Y., Funatsu G.;  
 RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";  
 RL Biochim. Biophys. Acta 872:277-285(1986).  
 RN [3]  
 RP SEQUENCE OF 303-337.  
 RX MEDLINE=80178723; PubMed=6768555;  
 RA Lin T.-S., Li S.-L.;  
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";  
 RL Eur. J. Biochem. 105:453-459(1980).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DB EMBL; M12089; AAA33869.1; -;  
 DB EMBL; S40368; AAB22584.1; -;

DR PIR; A24261; RLCSAG.  
 DR HSSP; P02879; LBR6.  
 DR GlycosuitedB; P06750;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS00231; RICIN B LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 290 AGGLUTININ A CHAIN.  
 FT PROPEP 291 302 LINKER PEPTIDE.  
 FT CHAIN 303 564 AGGLUTININ B CHAIN.  
 FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 319 361 1-ALPHA.  
 FT REPEAT 362 402 1-BETA.  
 FT REPEAT 405 437 1-GAMMA.  
 FT REPEAT 450 485 2-ALPHA.  
 FT REPEAT 489 528 2-BETA.  
 FT REPEAT 531 558 2-GAMMA.  
 FT ACT SITE 200 200 BY SIMILARITY.  
 FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 322 341 BY SIMILARITY.  
 FT DISULFID 365 382 BY SIMILARITY.  
 FT DISULFID 453 466 BY SIMILARITY.  
 FT DISULFID 492 509 BY SIMILARITY.  
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 331 331 F -> T (IN REF. 2).  
 FT CONFLICT 362 362 N -> D (IN REF. 2).  
 FT CONFLICT 374 374 R -> G (IN REF. 2).  
 FT CONFLICT 404 404 R -> T (IN REF. 2).  
 FT CONFLICT 552 552 F -> V (IN REF. 2).  
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match 90.8%; Score 930.5; DB 1; Length 564;  
 Best Local Similarity 91.9%; Pred. No. 5.6e-79;  
 Matches 182; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILV 61  
 DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILV 84

QY 62 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 121  
 DB 85 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 144

QY 122 GNYDRLEQLAGNLRNLELNGPLNEEALISALYYSTGCTQPLTARSFIICQMISEAAR 181  
 DB 145 GNYDRLEQLAGNLRNLELNGPLNEEALISALYYSTGCTQPLTARSFIICQMISEAAR 203

QY 182 FOYIEGEMRTRIRNRRS 199  
 DB 204 FOYIEGEMRTRIRNRRS 221

RESULT 3  
 ABRC ABRPR STANDARD; PRT; 562 AA.  
 AC P28590;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abrin-c precursor [Contains: Abrin-c A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-c B chain].  
 DE Abrus precatorius (Indian licorice) (Crab's eye).  
 OS

Best Local Similarity    45.5%;    Pred.No. 2.7e-25;  
Matches      86;    Conservative    24;    Mismatches    70;    Indels                  9;    Gaps                 4;

QY            6   QYEIINFTTAGATVQSYYTFIRAVRGLRTGADVRIEIPVLNVRVGLPINQRFILVELSN    65  
                 : | ||| ||||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :  
Db            35   QQDVIKFTEGATSQSKQFIKALRQLTGG--LIHDIPVPDPPTTVVEENRYITIVELSN    92

QY            66   HAEISVTILDVNAVVGVRAGNSAYFFHPDNQDEAAETHLFTDVCNYTFAFGGYD    125  
                 : | : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :  
Db            93   SPRESIEVDGVNAVVAIVRAGSQSYFL---RDAPASASTYLFPGTQ-RYSLFDFGSYG    148

QY            126   RIHQLAGNLRENIETNGPLEEIAISLYYSTQTGLPTLARSLFIICIMISEAARFOYI    185  
                 : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :  
Db            149   DIERWAHQTREEISLGLOALTHAIS---FLRS GASND EKARTLVIIOMASEAARYI    205

QY            186   EGEEMTRIR    194  
                 : |||  
Db            206   SNRVGVSVIR    214

RESULT 4

ABRR ABPR STANDARD; PRT; 527 AA.

ID \_ABRR ABPR ID \_ABRR AC Q06077; P81374; DT 15-DEC-1998 (Rel. 37, Created) DT 15-DEC-1998 (Rel. 37, Last sequence update) DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Abrin-b precursor [Contains: Abrin-b A chain (rRNA N-glycosidase) DE (EC 3.2.2.2); Abrin-b B chain]  
OS Abus precatorius (Indian licorice) (Crab's eye).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Abruceae; Abrus.  
ON NCBI\_TaxID=3816;  
RX [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=93132739; PubMed=8421313;  
RX Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;  
RT "Primary structure of three distinct isoabrins determined by cDNA sequencing. Conservation and significance."  
J. Mol. Biol. 229:263-267(1993).  
RN [2]  
RP SEQUENCE OF 260-527.  
RC TISSUE=Seed;  
RM MEDLINE=93169023; Pubmed=7763422;  
RA Kimura M., Sumizawa T., Funatsu G.;  
RT "The complete amino acid sequences of the B-chains of abrin-a and abrin-b, toxic proteins from the seeds of Abrus precatorius.";  
RL Biosci. Biotechnol. Biochem. 57:166-169(1993).  
CC -! FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
CC ABRIN-A IS MORE TOXIC THAN RICIN.  
CC -! FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
CC -! CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
CC -! SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
CC -! DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
CC -! SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
CC -! SIMILARITY: Contains 2 ricin B-type lectin domains.

-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----



```

DR EMBL; M98345; AAA32625.1; -.
DR PIR; S32430; S32430.
DR HSSP; P11140; LABR.
DR InterPro; IPR000772; Ricin_B_lectin.
DR DR InterPro; IPR001574; RIP.
DR DR Pfam; PF00652; Ricin_B_lectin; 6.
DR DR Pfam; PF00161; RIP; 1.
DR DR PRINTS; PM00396; SHIGARICIN.
DR DR SMART; SM00458; RICIN; 2.
DR DR PROSITE; PS0231; RICIN B LECTIN; 2.
DR DR PROSITE; PS0275; SHIGA RICIN; 1.
KW Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Pyrrolidone carboxylic acid.
FT CHAIN 1 250
FT PEPTIDE 251 260
FT ABRIN-B A CHAIN.
FT CHAIN 261 527
FT ABRIN-B B CHAIN.
FT DOMAIN 272 399
FT RICIN B-TYPE LECTIN 1.
FT DOMAIN 402 526
FT RICIN B-TYPE LECTIN 2.
FT REPEAT 282 324
FT 1-ALPHA.
FT REPEAT 325 365
FT 1-BETA.
FT REPEAT 368 400
FT 1-GAMMA.
FT REPEAT 413 448
FT 2-ALPHA.
FT REPEAT 452 491
FT 2-BETA.
FT REPEAT 494 527
FT 2-GAMMA.
FT ACT_SITE 163 163
FT BY SIMILARITY.
FT DISULFID 246 268
FT INTERCHAIN (BY SIMILARITY).
FT DISULFID 285 304
FT BY SIMILARITY.
FT DISULFID 328 345
FT BY SIMILARITY.
FT DISULFID 416 429
FT BY SIMILARITY.
FT DISULFID 445 472
FT BY SIMILARITY.
FT MOD_RES 1 1
FT PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT CARBOHYD 110 110
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 282 282
FT N -> D (IN REF. 2).
FT CONFLICT 291 291
FT D -> N (IN REF. 2).
FT CONFLICT 350 351
FT AE -> PO (IN REF. 2).
FT CONFLICT 378 378
FT S -> N (IN REF. 2).
FT CONFLICT 426 426
FT L -> M (IN REF. 2).
FT CONFLICT 428 428
FT Y -> D (IN REF. 2).
FT CONFLICT 431 431
FT N -> S (IN REF. 2).
FT CONFLICT 484 484
FT R -> K (IN REF. 2).
FT CONFLICT 491 491
FT N -> S (IN REF. 2).
FT CONFLICT 493 493
FT H -> Y (IN REF. 2).
FT CONFLICT 502 502
FT R -> G (IN REF. 2).
FT CONFLICT 509 509
FT E -> Q (IN REF. 2).
FT CONFLICT 513 513
FT H -> W (IN REF. 2).
FT CONFLICT 516 516
FT H -> T (IN REF. 2).
SQ SEQUENCE 527 AA; 59114 MW; 3253AE490CE9494A CRC64;

Query Match 33.7%; Score 345; DB 1; Length 527;
Best Local Similarity 45.5%; Pred. No. 1.5e-24;
Matches 87; Conservative 22; Mismatches 72; Indels 10; Gaps 4;

QY 6 QYPIINFTTAGTQSYNTFIRAVGRGLTGADVRHEIPVLPNRCVLPINQFVLVLSN 65
Db 1 QDQIKFTTEGATSQSYKQFIEALQRURGG--LHGIPVLPDPPTLQERNYIIVLSN 58

QY 66 HAELSVTLALDVNTAYVVGVRAGNSAYFPHPNQEDAEATHLFTDVQNRVTFAPFGNYD 125
Db 59 SDTESIEAGIDVSNAYVAVRAGNSYFL--RDAPTSASRYLFTCTQ-QYSLRFGSYI 114

QY 126 RLEQIAGNIRENIEUNGPLEIAISALYYSTGGTGLPTLARSFICIMISEARFQVI 185
Db 115 DLERLARQTRQIPLGLALRHAISFL----QSGTDDQEIARTLIVIIQWASEAARYREI 170

QY 186 EGEEMTRIRYN 196
Db 171 SYRVGVGSIRTN 181

RESULT 5

```

```

RIPT_TRIKI
ID RIPT_TRIKI STANDARD; PRT; 289 AA.
AC P09989;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein alpha-trichosanthin precursor
DE (RNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz;
RA MEDLINE=91153657; PubMed=199291;
RX Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia
RL coli.";
RL Gene 97:267-272 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz; TISSUE=Leaf;
RA MEDLINE=90256790; PubMed=2341400;
RX Chow T., Feldman R.A., Lovett M., Piatak M.;
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a
RL type I ribosome-inactivating protein.";
RL J. Biol. Chem. 265:8670-8674 (1990).
RN [3]
RP SEQUENCE OF 24-270.
RC STRAIN=Maximowicz; TISSUE=Tuberous root;
RA MEDLINE=90256789; PubMed=2341399;
RX Collins E.J., Robertus J.D., Ippresci M., Stone K.L., Williams K.R.,
RA Wu P., Hwang K., Piatak M.;
RT "Primary amino acid sequence of alpha-trichosanthin and molecular
RL models for abrin A-chain and alpha-trichosanthin.";
RL J. Biol. Chem. 265:8665-8669 (1990).
RN [4]
RP SEQUENCE OF 24-270.
RC TISSUE=Tuberous root;
RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,
RA Tian G.Y., Ni C.Z.;
RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and
RL application.";
RL Pure Appl. Chem. 58:789-798 (1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RA MEDLINE=94344957; PubMed=8066085;
RX Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT "Structure of trichosanthin at 1.88-A resolution.";
RL Proteins 19:4-13 (1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA MEDLINE=95344383; PubMed=7619070;
RX Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RL depurinating mechanism of two ribosome-inactivating proteins.";
RL Biochem. J. 309:285-298 (1995).
CC -!- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT
CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC CC EMBL; M34858; AAA34207.1; -;  
 DR DR EMBL; J05434; AAA34206.1; -;  
 DR DR PIR; J0566; RUTZ.  
 DR DR PDB; IMRJ; 07-FEB-95.  
 DR DR PDB; IMRJ; 07-FEB-95.  
 DR DR PDB; 1TCS; 10-JUL-95.  
 DR DR PDB; 1J4G; 28-JAN-03.  
 DR DR PDB; 1NLI; 21-JAN-03.  
 DR DR PDB; 1QD2; 24-APR-00.  
 DR DR InterPro; IPR001574; RIP.  
 DR DR Pfam; PF00161; RIP; 1.  
 DR DR PRINTS; PR00396; SHIGARICIN.  
 DR DR PROSITE; PS00275; SHIGA RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 270  
 FT FT 271 289  
 FT FT MISSING IN MATURE PROTEIN.  
 FT FT ACT SITE 183 183  
 FT FT CONFLICT 57 60  
 FT FT CONFLICT 82 84  
 FT FT CONFLICT 87 87  
 FT FT CONFLICT 92 92  
 FT FT CONFLICT 92 92  
 FT FT CONFLICT 143 144  
 FT FT CONFLICT 196 196  
 FT FT CONFLICT 215 216  
 FT FT CONFLICT 231 231  
 FT FT CONFLICT 234 234  
 FT FT CONFLICT 246 266  
 FT FT CONFLICT 247 247  
 FT FT STRAND 25 28  
 FT FT TURN 30 31  
 FT FT HELIX 34 46  
 FT FT TURN 47 47  
 FT FT STRAND 50 54  
 FT FT TURN 55 56  
 FT FT STRAND 57 60  
 FT FT HELIX 66 69  
 FT FT STRAND 70 76  
 FT FT TURN 78 79  
 FT FT STRAND 82 88  
 FT FT TURN 89 92  
 FT FT STRAND 93 99  
 FT FT TURN 100 101  
 FT FT STRAND 102 105  
 FT FT HELIX 109 114  
 FT FT TURN 115 117  
 FT FT TURN 120 121  
 FT FT STRAND 124 127  
 FT FT HELIX 134 141  
 FT FT TURN 142 142  
 FT FT HELIX 145 147  
 FT FT STRAND 150 150  
 FT FT HELIX 152 163  
 FT FT TURN 164 165  
 FT FT HELIX 167 180  
 FT FT TURN 181 181  
 FT FT HELIX 182 186  
 FT FT STRAND 187 187  
 FT FT HELIX 188 195  
 FT FT TURN 196 196  
 FT FT STRAND 202 202  
 FT FT HELIX 206 226  
 FT FT TURN 227 230  
 FT FT STRAND 231 239  
 FT FT TURN 241 242  
 FT FT STRAND 245 250  
 FT FT TURN 251 252  
 FT FT HELIX 254 258  
 FT FT TURN 259 259

FT STRAND 260 260  
 FT STRAND 263 263  
 FT TURN 266 268  
 SQ SEQUENCE 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;  
 Query Match 33.4%; Score 342; DB 1; Length 289;  
 Best Local Similarity 39.5%; Pred. No. 1.4e-24;  
 Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
 QY 10 INFTTAGATVQSYTNFIRAVRGRLTTGADVRHEIPVLPENRVGLPINQRFILVELSNHAEI 69  
 DB 25 VSFRLSGATSSSYGVFISNLKALPNERKL-YDIPLL--RSLPGSQRYALIHILNYADE 81  
 QY 70 SVTLADVTNAYVGYRAGNSAYFPHDPQEDA-EAITHLFTDVQNRVTFAGGNYDRLE 128  
 DB 82 TTSVAIDVTNVYINGRAGDTSYFF---NEASATEAAKYVFKDAMRKVTLTPYSGNYERIQ 138  
 QY 129 QLAGNLRNIELGNGPLBEAISALYYSTYGTGQTLPALRSFICICIMISEARFQYIEGE 188  
 DB 139 TAAGKIRENIPGLPALDSAITLIFYNAN-----SAASALMVLQSTSEAAKYKFEQQ 193  
 QY 189 METRI 193  
 DB 194 ICKRV 198  
 RESULT 6  
 RIP2\_BRYDI STANDARD; PRT; 282 AA.  
 AC P8184; Q988J0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).  
 OS Bryonia dioica (Red bryony).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 NCBI\_TaxID=3652;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA Siegal C.B., Gawlak S.L., Marquardt H.;  
 RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant  
 RL Bryonia dioica.";  
 RL Patent number US5597569, 28-JAN-1997.  
 [2]  
 RN SEQUENCE OF 22-42.  
 RC TISSUE=Root;  
 RX MEDLINE=95151812; PubMed=7849072;  
 RA Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
 RA Marquardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from  
 RT Bryonia dioica and their utility as carcinoma-reactive  
 RT immunoconjugates.";  
 RL Bioconj. Chem. 5:423-429 (1994).  
 CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; I34238; -; NOT\_ANNOTATED\_CDS.  
 DR HSP; P09989; 1MRJ.

```
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA RICIN; 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 282
FT ACT_SITE 183 183
FT CARBOHYD 25 25
SQ SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;

Query Match 33.3%; Score 341.5; DB 1; Length 282;
Best Local Similarity 48.0%; Pred. No. 1.5e-24;
Matches 86; Conservative 23; Mismatches 57; Indels 13; Gaps 7;

QY 10 INFTTAGATVQSYNFIRAVRGRLTTGADVREHPVLPN-RVGLPIINQRFILVELSNHAE 68
Db 24 INFTSLGATGATYKTFIRNLRKLTGTGPRVDIPVLRNAAAGL--ARQLVLTNYNG 80
QY 69 LSVTLALDVTNAVYVGYRAGNSAYFFPHDQEDAEATHLFTDVQNYRTFAFGNYDRLE 128
Db 81 ESVTVALDVNVVYVYRAGNTAYFL--ADASTEANNVLFAGI-NHVRLPYGGNYDGL 136
QY 129 QLAGNL-RENIELGNGPLEEASALYYSTGGTQPLTARSFFIICMISEAAARFQYIE 186
Db 137 TAAGRISRENIELGFSEISSAIGNMFRHNP-GTSVP---RAFIVIIQTVSEAAARFQYIE 191

RESULT 7
RIPS TRIKI
ID -RIPS TRIKI STANDARD; PRT; 289 AA.
AC P24478;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein karasurin precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Root tuber;
RX MEDLINE=97356562; PubMed=9212998;
RA Mizukami H., Iida K., Kondo T., Ogihara Y.;
RT "Cloning and bacterial expression of a gene encoding ribosome-
RT inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
RT kirilowii var. japonica.";
RL Biol. Pharm. Bull. 20:711-713 (1997).
[2]
RN SEQUENCE OF 24-270.
RP MEDLINE=92005921; PubMed=1914000;
RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,
RT karasurin.";
RL Chem. Pharm. Bull. 39:1244-1249 (1991).
CC -!- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC
CC 60S RIBOSOMAL SUBUNITS.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
CC EMBL; AB000666; BAA21786.1; -.
CC PIR; JCS606; JCS606.
CC PIR; JU0393; JU0393.
CC HSSP; P09989; LMRJ.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC PROSITE; PS00275; SHIGA RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal.
FT SIGNAL 1 21
FT CHAIN 22 270
FT CHAIN 24 270
FT PROPEP 271 289
FT ACT_SITE 183 183
SQ SEQUENCE 289 AA; 31704 MW; 883D3E3242887B26 CRC64;

Query Match 33.0%; Score 338; DB 1; Length 289;
Best Local Similarity 40.0%; Pred. No. 3.3e-24;
Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5;

QY 10 INFTTAGATVQSYNFIRAVRGRLTTGADVREHPVLPN-RVGLPIINQRFILVELSNHAE 69
Db 25 VSRFSGATSSSYGVFISNLRKALPYERKL-YDIPLL--RSTLPGSORYALIHLYVADE 81
QY 70 SVTLALDVTNAVYVGYRAGNSAYFFPHDQEDA-EATHLFTDVQNYRTFAFGNYDRLE 128
Db 82 TISVAIDVTNVYVGYRAGDTSYFF--NEASATEAAKYVFKDKRKVTLPYSGNYERLQ 138
QY 129 QLAGNLRENIELGNGPLEEASALYYSTGGTQPLTARSFFIICMISEAAARFQYIEGE 188
Db 139 IAAGKIRENIPLGLDPLDLSAITFLFYNN-----SAASALMVLIQSTSEAAKYKFIEQ 193
QY 189 MRTRI 193
Db 194 IGKRV 198

RESULT 8
NIGB SAMNI
ID -NIGB SAMNI STANDARD; PRT; 563 AA.
AC P33183; P33184; P93542;
DT 01-OCT-1993 (Rel. 27, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nigrin b precursor (Agglutinin V) (SNAV) [Contains: Nigrin b A chain
DE (rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain].
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.
OX NCBI_TaxID=4202;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Bark;
RX MEDLINE=96215449; PubMed=8647092;
RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;
RT "Characterization and molecular cloning of Sambucus nigra agglutinin V
RT (nigrin b), a GalNAc-specific type-2 ribosome-inactivating protein
RT from the bark of elderberry (Sambucus nigra).";
RL Eur. J. Biochem. 237:505-513 (1996).
[2]
RN SEQUENCE OF 26-49 AND 298-321.
RP TISSUE=Bark;
RX MEDLINE=94003077; PubMed=8400135;
RA Girbes T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R.,
RA Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;
RT "Isolation and partial characterization of nigrin b, a non-toxic
RT novel type 2 ribosome-inactivating protein from the bark of Sambucus
RT nigra L.";
RL Eur. J. Biochem. 237:505-513 (1996).
[2]
RN SEQUENCE OF 26-49 AND 298-321.
RP TISSUE=Bark;
RX MEDLINE=94003077; PubMed=8400135;
RA Girbes T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R.,
RA Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;
RT "Isolation and partial characterization of nigrin b, a non-toxic
RT novel type 2 ribosome-inactivating protein from the bark of Sambucus
RT nigra L.";
RL Plant Mol. Biol. 22:1181-1186 (1993).
CC -!- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN
```



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M98344; AAA32624.1; ALT INIT.  
 DR EMBL; X54872; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; S32429; TZLSA.  
 DR PDB; 1ABR; 07-FEB-95.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; P00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS0275; SHIGA\_RICIN; 1.  
 KW Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.  
 FT CHAIN 1 251  
 FT PEPTIDE 252 261  
 FT CHAIN 262 528  
 FT DOMAIN 273 400  
 FT DOMAIN 403 527  
 FT REPEAT 283 325  
 FT REPEAT 326 366  
 FT REPEAT 369 401  
 FT REPEAT 414 449  
 FT REPEAT 453 492  
 FT REPEAT 495 528  
 FT ACT SITE 164 164  
 FT DISULFID 247 269  
 FT DISULFID 286 305  
 FT DISULFID 329 346  
 FT DISULFID 417 430  
 FT DISULFID 456 473  
 FT MOD RES 1 1  
 FT CARBOHYD 361 361  
 FT CARBOHYD 401 401  
 FT CONFLICT 202 202  
 FT CONFLICT 298 298  
 FT CONFLICT 427 427  
 FT CONFLICT 457 457  
 FT CONFLICT 483 483  
 FT STRAND 5 8  
 FT TURN 10 11  
 FT HELIX 14 28  
 FT STRAND 32 33  
 FT TURN 34 35  
 FT STRAND 36 38  
 FT TURN 42 43  
 FT HELIX 47 49  
 FT STRAND 51 57  
 FT TURN 63 69  
 FT TURN 70 72  
 FT STRAND 75 79  
 FT STRAND 83 86  
 FT TURN 88 89  
 FT TURN 92 93  
 FT HELIX 94 97  
 FT TURN 100 101  
 FT STRAND 103 106  
 FT TURN 113 114  
 FT HELIX 115 119  
 FT TURN 124 126  
 FT STRAND 129 129  
 FT HELIX 131 142  
 FT TURN 143 144  
 FT HELIX 148 167  
 FT STRAND 168 168  
 FT HELIX 169 180  
 FT TURN 181 182

FT STRAND 185 185  
 FT HELIX 189 196  
 FT TURN 197 207  
 FT HELIX 198 207  
 FT STRAND 212 220  
 FT TURN 222 223  
 FT STRAND 226 231  
 FT TURN 232 233  
 FT HELIX 235 239  
 FT STRAND 240 240  
 FT STRAND 243 243  
 FT STRAND 248 248  
 FT STRAND 268 268  
 FT STRAND 276 277  
 FT STRAND 279 280  
 FT HELIX 282 284  
 FT STRAND 286 289  
 FT HELIX 290 292  
 FT TURN 296 297  
 FT STRAND 299 303  
 FT HELIX 311 313  
 FT STRAND 315 317  
 FT TURN 320 320  
 FT STRAND 322 325  
 FT TURN 326 327  
 FT STRAND 328 332  
 FT TURN 337 338  
 FT STRAND 340 344  
 FT TURN 346 348  
 FT HELIX 351 353  
 FT STRAND 355 355  
 FT STRAND 357 358  
 FT TURN 360 361  
 FT STRAND 364 366  
 FT TURN 371 374  
 FT STRAND 371 374  
 FT TURN 380 381  
 FT STRAND 383 383  
 FT STRAND 385 387  
 FT HELIX 393 395  
 FT STRAND 398 399  
 FT STRAND 406 408  
 FT STRAND 410 411  
 FT HELIX 413 415  
 FT STRAND 417 421  
 FT TURN 422 423  
 FT STRAND 424 428  
 FT TURN 432 433  
 FT HELIX 435 437  
 FT STRAND 439 441  
 FT TURN 443 444  
 FT STRAND 447 449  
 FT TURN 450 451  
 FT STRAND 452 459  
 FT TURN 464 465  
 FT STRAND 467 472  
 FT TURN 474 475  
 FT HELIX 478 480  
 FT STRAND 483 484  
 FT TURN 486 487  
 FT STRAND 490 492  
 FT TURN 493 496  
 FT STRAND 497 501

Query Match 32.1%; Score 329.5; DB 1; Length 528;  
 Best Local Similarity 42.8%; Pred. No. 4.2e-23;

Matches 80; Conservative 28; Mismatches 66; Indels 13; Gaps 5;

QY 10 INFTTAGATVQSVTFIRAVGRGLTTGADVREHPVLPNRRVGLPINQRFLVLSNHAEL 69

Db 5 IKFTGATGSQSKYKQFIEALRELRGG--LIHIDIPVLPDPTTLQERNRITVLSNSDTE 62

QY 70 SVTLALDVTNAYVVGVRAGNSAYFFH--PDNQEDAEAIHTLFTDVQNRVYTFAGGNYDL 127

Db 63 STEVGIDVTNAYVAVRAGTQSFLEADPSASD-----YLFGT-DQSLPFGYGD 116  
QY 128 EQAGNLRENELNGPFLAEIAISALYIYSTGQTFLARSLFICIQMISEAARFOYIEG 187  
Db 117 ERWAHQSRQIPLGLQALTHGIS--FFRSGGNDNEKARTLVIQWAEARFRYISN 173  
QY 188 EMETRIR 194  
Db 174 RVRVSIQ 180

## RESULT 10

RIP1\_BRYDI  
ID RIP1\_BRYDI STANDARD; PRT; 290 AA.  
AC P3185; Q9S819;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)  
(EC 3.2.2.22) (BD1).  
OS Bryonia dioica (Red Bryony).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
OX NCBI\_TaxID=3652;  
RN [1]  
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RC TISSUE=Leaf;  
RX MEDLINE=97228081; PubMed=9115985;  
RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,  
RA Siegall C.B.,  
RT "Molecular, biological, and preliminary structural analysis of  
RT recombinant bryodin I, a ribosome-inactivating protein from the plant  
RT Bryonia dioica."  
RL Biochemistry 36:3095-3103 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Siegall C.B.;  
RT "Cloning and expression of a gene encoding bryodin I from Bryonia  
RT dioica."  
RL Patent number US5541110, 30-JUL-1996.  
RN [3]  
RP SEQUENCE OF 24-66.  
RC TISSUE=Seed;  
RX MEDLINE=89326691; PubMed=2753596;  
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,  
RA Lappi D.;  
RT "N-terminal sequence of some ribosome-inactivating proteins."  
RL Int. J. Pept. Protein Res. 33:263-267 (1989).  
RN [4]  
RP SEQUENCE OF 24-43.  
RC TISSUE=Root;  
RX MEDLINE=95151812; PubMed=7849072;  
RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
RA Marquardt H.;  
RT "Characterization of ribosome-inactivating proteins isolated from  
RT Bryonia dioica and their utility as carcinoma-reactive  
RT immunoconjugates."  
RL Bioconj. Chem. 5:423-429 (1994).  
CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
CC PROTEIN SYNTHESIS IN ANIMAL CELLS.  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S rRNA.  
CC -!- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO  
CC PRODUCE A SHORTER PROTEIN.  
CC -!- BIOTECHNOLOGY: Especially useful as immunotoxin for  
CC pharmacological applications as it has low toxicity in rats and  
CC mice but is potent once inside target cells  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC TYPE 1 RIP SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; I24020; -; NOT\_ANNOTATED\_CDS.  
DR FIR; S18491; S16491.  
DR PDB; 1BRY; 04-MAR-98.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
KW 3D-structure; Multigene family; Glycoprotein; Signal.  
FT SIGNAL 1 23  
FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.  
FT PROPEP 271 290 MISSING IN NATURE PROTEIN.  
FT ACT\_SITE 183 183 BY SIMILARITY.  
FT ACT\_SITE 212 212  
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT MUTAGEN 212 212 E->K; REDUCES ACTIVITY 10-FOLD.  
FT CONFLICT 61 65 RSSIS -> LRHXI (IN REF. 3).  
FT STRAND 25 28  
FT TURN 30 31  
FT HELIX 34 46  
FT TURN 47 47  
FT STRAND 50 54  
FT TURN 55 56  
FT STRAND 57 60  
FT HELIX 66 69  
FT STRAND 70 76  
FT TURN 78 79  
FT STRAND 82 88  
FT TURN 89 92  
FT STRAND 93 99  
FT TURN 100 101  
FT STRAND 102 105  
FT HELIX 109 114  
FT TURN 115 117  
FT TURN 120 121  
FT STRAND 124 127  
FT HELIX 134 141  
FT TURN 142 142  
FT HELIX 145 147  
FT STRAND 150 150  
FT HELIX 152 163  
FT TURN 164 165  
FT HELIX 167 186  
FT STRAND 187 187  
FT HELIX 188 196  
FT STRAND 202 202  
FT HELIX 206 213  
FT TURN 214 214  
FT HELIX 215 225  
FT TURN 226 230  
FT STRAND 231 239  
FT TURN 241 242  
FT STRAND 245 250  
FT TURN 251 252  
FT HELIX 254 257  
FT TURN 258 259  
FT STRAND 260 260  
FT STRAND 263 263  
FT HELIX 266 268  
SQ SEQUENCE 290 AA; E966CD9C031A42DB CRC64;

Query Match 30.4%; Score 312; DB 1; Length 290;  
Best local Similarity 37.3%; Pred. No. 8.6e-22;  
Matches 69; Conservative 46; Mismatches 58; Indels 12; Gaps 5;

QY 10 INFTAGATVQSYTNFIRAVRGRLLTTGADVRHEIPVLNPRVGLPINORFILVELSNHAEL 69  
 Db 25 VSFRLSGATTTSYGVFKNLREALPYERKV-YNIPLL--RSSISGSGRYTLHLHTNYADE 81  
 QY 70 SVTLALDVTNAYVVGVRAGNSAYFFHPDQORDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 Db 82 TISVADVDTNVIYINGVLADVSYFF--NEASATEAAKFVKDAKKVTLPLPSGNYERLQ 138  
 QY 129 QLGNLRENTLGNPLLEEAISALYYSTGTQTLPLARSFIICIMISEARFQYIEGE 188  
 Db 139 TAAGKIRENPLGLPALDSAITLLYYVTAS-----SAASALLVLIQSTAESARYKFFIEQ 193  
 QY 189 MTRPI 193  
 Db 194 IGRKV 198

## RESULT 11

RIPB LUFYC  
 ID RIPB LUFYC STANDARD; PRT; 250 AA.  
 AC P22851;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein luffin-B (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 OS Luffa cylindrica (Smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 OX NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Seed;  
 RX MEDLINE=91248488; PubMed=1368666;  
 RA Islam M.R., Hirayama H., Funatsu G.;  
 RT "Complete amino acid sequence of luffin-b, a ribosome-inactivating  
 RT protein from sponge gourd (Luffa cylindrica) seeds.";  
 RL Agric. Biol. Chem. 55:229-238(1991).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 DR PIR; JN0108; JN0108.  
 DR HSP; P16094; IAH.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin.  
 FT ACT SITE 160 160 BY SIMILARITY.  
 SQ SEQUENCE 250 AA; 27293 MW; FOIA8DC8A107800 CRC64;

Query Match 29.3%; Score 300.5; DB 1; Length 250;  
 Best Local Similarity 34.2%; Pred. No. 8.4e-21;  
 Matches 64; Conservative 48; Mismatches 64; Indels 11; Gaps 3;

QY 10 INFTAGATVQSYTNFIRAVRGRLLTTGADVRHEIPVLNPRVGLPINORFILVELSNHAEL 69  
 Db 3 VSFRLSGATTSYGVFKNLREALPYERKV-YNIPLL--RSSISGSGRYTLHLHTNYADE 81  
 QY 70 SVTLALDVTNAYVVGVRAGNSAYFFHPDQORDA-EAITHLFTDVQNRVYTFAGGNYDRLEQ 129  
 Db 60 AITMAIDVTNVIYINGVLADVSYFF--ANESAKLASQVFKGSLVTPISGNYERLQN 116  
 QY 130 LAGNIRENTLGNPLLEEAISALYYSTGTQTLPLARSFIICIMISEARFQYIEGEM 189  
 Db 117 AAGKIRENPLGLPALDSAITLLYYVTAS-----TAAAFVILQTTAESRKFYIEQI 171  
 QY 190 MTRPI 196  
 Db 172 IERIPN 178

RESULT 12  
 RIP1 CUCFI  
 ID RIP1 CUCFI STANDARD; PRT; 286 AA.  
 AC Q9PRX4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22).  
 OS Cucumis figarei.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=131071;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada T., Ohki S.T., Osaki T.;  
 RT "Cloning and analysis of a cDNA coding a putative ribosome-  
 RT inactivating protein from Cucumis figarei.";  
 RL Plant Biotechnol. 17:337-340(2000).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC EMBL; AB045560; BAB19677.1; -  
 DR HSP; P16094; IAH.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 286 PUTATIVE RIBOSOME-INACTIVATING PROTEIN.  
 FT ACT SITE 185 185 BY SIMILARITY.  
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 286 AA; 31771 MW; 4EFD4966E604DA41 CRC64;

Query Match 28.9%; Score 296.5; DB 1; Length 286;  
 Best Local Similarity 34.4%; Pred. No. 2.3e-20;  
 Matches 65; Conservative 47; Mismatches 64; Indels 13; Gaps 5;

QY 10 INFTAGATVQSYTNFIRAVRGRLLTTGADVRHEIPVLNPRVGLPINORFILVELSNHAEL 68  
 Db 28 VKFSLGNSNHSKSFITSMRNALPNAGDI-YNIPLLVPSISG---SRVILMQLSNYEG 83  
 QY 69 LSVTLALDVTNAYVVGVRAGNSAYFFHPDQORDA-EAITHLFTDVQNRVYTFAGGNYDRLE 128  
 Db 84 NTITMAIDVTNVIYINGVLADVSYFF--NETDAQASKFVFGTKSITLPSYGNKYQLQ 140  
 QY 129 QLGNLRENTLGNPLLEEAISALYYSTGTQTLPLARSFIICIMISEARFQYIEGE 188  
 Db 141 SVAREKRDSPGLGPMALDSAITLLYYDSRAPI-----AFLVLIQTAAEAARYKIEKQ 195  
 QY 189 MTRPI 197  
 Db 196 IIDRISVK 204

## RESULT 13

MLA\_VISAL ID MLN\_VISAL STANDARD; PRT; 254 AA.

AC P81446;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-glycosidase) (EC 3.2.2.22).  
 OS Viscum album (European mistletoe).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Santalales; Santalaceae; Viscum.  
 OX NCBI\_TaxID=3972;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=Subsp. album;  
 RX MEDLINE=97134581; PubMed=8980141;  
 RA Soler M.H.; Stoeva S.; Schwaborn C.; Wilhelm S.; Stiefel T.;  
 RA Voelter W.;  
 RT "Complete amino acid sequence of the A chain of mistletoe lectin I."; FEBS Lett. 399:153-157(1996).  
 RL FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S rRNA. THE B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY INTO THE CELL OF THE A CHAIN; B CHAINS ARE ALSO RESPONSIBLE FOR CELL AGGLUTINATION (LECTIN ACTIVITY).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.  
 CC -1- PHARMACOLOGICAL: Due to its immunomodulative effects it is being studied in clinical trials in cancer patients as it may slow the growth of cancer cells and be an effective treatment for solid tumors.  
 CC -1- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA AND NON-GLYCOSYLATED FORM MLA'.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 2 RIP SUBFAMILY.  
 DR PIR; PD0018; PD0018.  
 DR HSP; P11140; IABR.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; FALSE NEG.  
 KW Plant defense; Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.  
 FT ACT\_SITE 165 165  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC...)  
 FT VARIANT 15 15 E -> D (IN MLA').  
 FT VARIANT 66 66 V -> I (IN MLA').  
 FT VARIANT 112 112 N -> T (IN MLA').  
 FT VARIANT 116 116 P -> T (IN MLA').  
 FT VARIANT 133 134 DQ -> EE (IN MLA').  
 FT VARIANT 140 140 T -> S (IN MLA').  
 FT VARIANT 144 144 F -> Y (IN MLA').  
 FT VARIANT 151 151 T -> A (IN MLA').  
 FT VARIANT 179 179 Y -> D (IN MLA').  
 FT VARIANT 184 184 A -> E (IN MLA').  
 FT VARIANT 190 190 V -> M (IN MLA').  
 FT VARIANT 218 218 I -> F (IN MLA').  
 FT VARIANT 223 224 PP -> ST (IN MLA').  
 FT VARIANT 231 231 T -> S (IN MLA').  
 FT VARIANT 235 235 D -> S (IN MLA').  
 FT SEQUENCE 254 AA; 28478 MW; 53BAF98D3E0FFE67 CRC64;  
 Query Match 27.9%; Score 286; DB 1; Length 254;  
 Best Local Similarity 38.4%; Pred. No. 1.9e-19;  
 Matches 81; Conservative 36; Mismatches 56; Indels 38; Gaps 10;  
 QY 14 TAGATVQSYTNFIRAVRGRLTTCADVRHEIPVLPNRVGLPILN--QRFILVELSNHAEISLV 71  
 Db 9 THQTGEEYFRFTLLRDYVSSGS--FSNEIPLL-RQSTIVSDRQRFVLVELTQQGQSV 66  
 QY 72 TLALDVNTNAYVVGVRAGNSAYFFHPDNCDEAEATHLFTDVQNRVYTFAGGNYDRLEQLA 131

Db 67 TAAIDVTNAYVVGVRAGNSAYFLR-DAPRGAE--THLFTGT-TRSLFPNGSYPLDERVA 122  
 QY 132 GNLENIELNGNPLEEATISALYYVYSTGTQPTLARSFLICMIOMISEAARF-----Q 183  
 Db 123 GH-RDQIPGLIDQITQSVTALRF---PGSGTRTQARSILILMIOMISEAARFNPILWRYRQ 178  
 QY 184 YIE-----GEMETRIYN 196  
 Db 179 YINGASFLPDVYMLELETSWGQQSTQVQHS 209  
 RESULT 14  
 ID RIPX\_LUFFCY STANDARD; PRT; 277 AA.  
 AC Q00465;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein luffin-alpha precursor (rRNA N-glycosidase) (EC 3.2.2.22).  
 OS Luffa cylindrica (Smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 OX NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed;  
 RX MEDLINE=9228316; PubMed=1600156;  
 RA Kataoka J.; Habuka N.; Miyano M.; Masuta C.; Koizumi A.;  
 RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-inactivating protein from Luffa cylindrica.";  
 RL Plant Mol. Biol. 18:1199-1202(1992).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X62371; CAA44229.1; -.  
 DR PIR; S22494; S22494.  
 DR HSP; P16094; LAHC.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.  
 FT SIGNAL 1 19 RIBOSOME-INACTIVATING PROTEIN LUFFIN-  
 FT CHAIN 20 277 ALPHA.  
 FT ACT\_SITE 179 179 BY SIMILARITY.  
 FT SEQUENCE 277 AA; 30212 MW; EAI7FC27998C25AC CRC64;  
 Query Match 27.8%; Score 284.5; DB 1; Length 277;  
 Best Local Similarity 33.5%; Pred. No. 2.9e-19;  
 Matches 63; Conservative 45; Mismatches 69; Indels 11; Gaps 3;  
 QY 10 INTTACATVQSYTNFIRAVRGRLTTCADVRHEIPVLPNRVGLPILNQRFLVELSNHAEI 69  
 Db 22 VRESLSSGSSSTSYKFTGDKALPNSGTVYNTLLISSASGA---SRYTILMTLSYDVGK 78  
 QY 70 SVTLALDVNTNAYVVGVRAGNSAYFFHPDNCDEAEATHLFTDVQNRVYTFAGGNYDRLEQ 129  
 Db 79 AITVADVNTVNYINGVLVNSTSYFF---NESDAKLASQYVFKGSTIVTLTPYSGNYEKLQT 135



QY		130	L A G N U R E N I E I G N P L E B A I S A L Y X Y T G G T Q L P T L A R S F I I C I O M I S E A A R Q Y I T E G E M . 189
	:		: : : : :
DB		136	A A K R I X E I P I G P A L D S A I T T L P H Y D S - - - - T A A A A A F L V I I Q T T A E A S R F K Y I E G O I . 190
	:		: : : : :
QY		190	R T R I R Y N R . 197
	:		: : :
DB		191	I E R I S K N Q . 198

RESIST 15

RESULT	IP23	MOMBA	STANDARD;	PRT;	286 AA.
AD	RIP2	MOMBA			
AC	P29339;				
DT	01-DEC-1992	(Rel. 24, Created)			
DT	01-DEC-1992	(Rel. 24, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Ribosome-inactivating protein momordin II precursor (rRNA				
DE	N-glycosidase) (EC 3.2.2.22).				
OS	Momordica balsamina (Bitter melon) (Balsam pear).				
OC	Eukaryota;	Viridiplantae;	Stertophyta;	Embryophyta;	Tracheophyta;
OC	Spermatophyta;	Magnoliophyta;	eudicotyledons;	core eudicots;	Rosidae;
OC	eurosid I;	Cucurbitales;	Cucurbitaceae;	Momordica.	
OX	NCBI TaxID=3672;				

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL:	Z12175;	CAA78166.1;	-.
DR	PIR:	S25560;	S25560.	
DR	PDB:	LCFS;	07-JUN-99.	
DR	InterPro:	IPR001574;	RIP.	
DR	Fram:	PF00161;	RIP; 1.	
DR	PRINTS;	PR00396;	SHIGARICIN.	
DR	PROSITE;	PS00275;	SHIGA.RICIN; 1.	
KW	Plant defense;	Protein synthesis inhibitor;	Hydrolase; Toxin; Signal;	
KW	3D-structure.			
FT	SIGNAL	1	23	
FT	CHAIN	24	286	
FT				RIBSOME-INACTIVATING PROTEIN MOMORDIN
FT				II.
FT	ACT SITE	181	181	BY SIMILARITY.
SQ	SEQUENCE	286 AA;	32031 MW;	3889FFLA6B25986 CRC64;

```

DB      | :|:|:| | | :|:|:| : : : | :|:|:|:| :
138 AAHKRENIDIGLPAISATITLFYNA-----QSAPSALLVLIQTAAARFKYIERHV 192

QY      190 RTRIRYN 196

DB      : |
193 AKYVATN 199

```

Search completed: February 10, 2004, 16:23:23  
Job time : 7.63825 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 26.553 Seconds  
(without alignments)  
1933.961 Million cell updates/sec

Title: US-10-083-336A-5  
Perfect score: 1025  
Sequence: 1 MIFPKQYPIINFTRAGATVQ.....ARFQVIEGEMTRIRYNRRS 199

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 23.\*

- 1: sp archaea.\*
- 2: sp bacteria.\*
- 3: sp fungi.\*
- 4: sp human.\*
- 5: sp invertebrate.\*
- 6: sp mammal.\*
- 7: sp mhc.\*
- 8: sp organelle.\*
- 9: sp phage.\*
- 10: sp plant.\*
- 11: sp rodent.\*
- 12: sp virus.\*
- 13: sp vertebrate.\*
- 14: sp unclassified.\*
- 15: sp rvirus.\*
- 16: sp bacteriaph.\*
- 17: sp archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1017	99.2	541	Q41174	ricinus com
2	401.5	39.2	580	Q94BW3	Q94BW3 cinnamomum
3	397.5	38.8	580	Q94BW4	Q94BW4 cinnamomum
4	397.5	38.8	581	Q94BW5	Q94BW5 cinnamomum
5	395.5	38.6	549	Q9FV22	Q9FV22 cinnamomum
6	395.5	34.2	563	Q04367	Q04367 sambucus ni
7	347.5	33.9	564	Q9AVR2	Q9AVR2 sambucus ni
8	344.5	33.6	528	Q06076	Q06076 abrus preca
9	340	33.2	289	Q94KE4	Q94KE4 trichosan
10	338	33.0	247	Q9LRE3	Q9LRE3 trichosan
11	338	33.0	289	Q41216	Q41216 trichosan
12	336.5	32.8	252	Q38760	Q38760 abrus preca
13	330.5	32.2	563	Q945S2	Q945S2 sambucus ni
14	330.5	32.2	563	Q8GT32	Q8GT32 sambucus ni
15	329.5	32.1	252	Q38761	Q38761 abrus preca
16	324	31.6	270	Q8LPV7	Q8LPV7 trichosan

17	323.5	31.6	251	10	Q96236	Q96236 abrus preca
18	322.5	31.5	251	10	Q96237	Q96237 abrus preca
19	319	31.1	565	10	Q04071	Q04071 sambucus ni
20	317	30.9	270	10	Q41611	Q41611 trichosan
21	316.5	30.9	251	10	Q96235	Q96235 abrus preca
22	314.5	30.7	547	10	Q9M6E9	Q9M6E9 abrus preca
23	310.5	30.3	566	10	Q04072	Q04072 sambucus ni
24	307.5	30.0	278	10	Q00980	Q00980 luffa cylin
25	291	28.4	570	10	Q41358	Q41358 sambucus ni
26	285	27.8	249	10	Q8LKQ5	Q8LKQ5 viscum albu
27	285	27.8	570	10	Q22415	Q22415 sambucus ni
28	283.5	27.7	604	10	Q9M654	Q9M654 polygonatum
29	283	27.6	251	10	Q8LKQ4	Q8LKQ4 viscum albu
30	282	27.5	254	10	Q8LKQ6	Q8LKQ6 viscum albu
31	280	27.3	264	10	Q9FSH2	Q9FSH2 momordica c
32	278	27.1	293	10	Q8S452	Q8S452 jatropa cu
33	277.5	27.1	565	10	Q8W243	Q8W243 viscum albu
34	277	27.0	286	10	Q9FUV7	Q9FUV7 momordica c
35	277	27.0	569	10	P93543	P93543 sambucus ni
36	275	26.8	531	10	Q8RXH6	Q8RXH6 viscum albu
37	273	26.6	249	10	Q8RXH7	Q8RXH7 viscum albu
38	272	26.5	286	10	Q41257	Q41257 momordica c
39	272	26.5	298	10	Q04358	Q04358 iris hollan
40	270	26.3	592	10	Q8W2E7	Q8W2E7 iris hollan
41	268	26.1	300	10	Q04356	Q04356 iris hollan
42	266.5	26.0	603	10	Q9M653	Q9M653 polygonatum
43	262.5	25.6	299	10	Q8GZM9	Q8GZM9 euphorbia s
44	261	25.5	293	10	Q8VYU0	Q8VYU0 jatropa cu
45	259	25.3	573	10	Q8W2E8	Q8W2E8 iris hollan

## ALIGNMENTS

RESULT 1

Q41174	PRELIMINARY;	PRT;	541 AA.
ID	Q41174		
AC	Q41174;		
DT	01-NOV-1996 (Tremblrel. 01, Created)		
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	Protein A chain (EC 3.2.2.22) (rRNA N-glycosidase)		
DE	(Fragment).		
OS	Ricinus communis (Castor bean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	euroside I; Malpighiales; Euphorbiaceae; Ricinus.		
OX	NCBI_TaxID=3988;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92338377; PubMed=1633311;		
RA	Roberts L.M., Tregear J.W., Lord J.M.;		
RT	"Molecular cloning of ricin."		
RL	Targeted Diagn. Ther. 7:81-97(1992).		
CC	-!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE		
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.		
CC	-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.		
DR	EMBL; S40366; AAB22582.1; -.		
DR	HSSP; P02879; 1BR6.		
DR	InterPro; IPR000772; Ricin_B_lectin.		
DR	InterPro; IPR001574; RIP.		
DR	InterPro; IPR001400; Somatotropin.		
DR	Pfam; PF00652; Ricin_B_lectin; 6.		
DR	Pfam; PF00161; RIP; 1.		
DR	PRINTS; PR00396; SHIGARICIN.		
DR	SMART; SM00458; RICIN; 2.		
DR	PROSITE; PS00231; RICIN_B_LECTIN; 2.		
DR	PROSITE; PS00275; SHIGA_RICIN; 1.		
DR	PROSITE; PS00338; SOMATOTROPIN_2; 1.		
KW	Hydrolase; Toxin.		
FT	NON TER		
SEQUENCE	541 AA; 60281 MW; 2B7B2CDEF1F2E9D9 CRC64;		

[illegible]

```

DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 5.
DR Pfam; PF00161; RIP; I_lectin.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS0231; RICIN_B_LECTIN; 2.
DR Hydrolase; 3.
FT NON TER 1
SQ SEQUENCE 549 AA; 60648 MW; 02607FE607CA44B0 CRC64;

Query Match 38.8%; Score 395.5; DB 10; Length 549;
Best Local Similarity 50.0%; Pred. No. 9.5e-29;
Matches 94; Conservative 27; Mismatches 60; Indels 7; Gaps 5;

QY 7 YPIINFTTAGATVQSYTNFIRAVRGRLITTCADVRHEIPVLNPNVGLPINORFILVELSN- 65
DB 1 YQIVTTTKKATKTSYTFEALRAQLASGEE-PHGIPVWRERSTVPDSKRFILVELSNW 59
QY 66 HAEISVTLALDVNTNAYVGVRAINSAYFFHPDNOEDAEATHLFTDVQNYRYTTFAGGNYD 125
DB 60 AADSPVTLAVDVTNAYVAVRTGQSFFLRDNP--PAIENLLPDTK-RYTFPFGSGYT 116
QY 126 RLRLQAGNRETELGNGLPEERALSALYYSTGGTQLPRLARSFIICIMSEARQYI 185
DB 117 DLEGVAGEREEILLGMPLLENALSALWISNL--NOORALARSLLIVVQWVAEVRPFI 174
QY 186 EGEEMRTRI 193
DB 175 EYRVGSI 182

RESULT 6
004367 PRELIMINARY; PRT; 563 AA.
AC
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Ribosome inactivating protein precursor (EC 3.2.2.22) (rRNA
DE N-glycosidase).
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.
NCBI_TaxID=4202;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112023; PubMed=9450339;
RA Van Damme E.J., Roy S., Barre A., Rouge P., Van Leuven F.,
RA Peumans W.J.;
RT "The major elderberry (Sambucus nigra) fruit protein is a lectin
RT derived from a truncated type 2 ribosome-inactivating protein.";
RL Plant J. 12:1251-1260(1997).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; U76524; AAC1586.1; -.
DR HSP; P02879; 2AAL.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; I_lectin.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS0231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA RICIN; 1.
DR Hydrolase; Signal; Toxin.
FT SIGNAL 1 25
FT CHAIN 26 297
FT CHAIN 298 563
SQ SEQUENCE 563 AA; 62336 MW; 3ED2B6C08E796205 CRC64;

```

Query Match 34.2%; Score 350.5; DB 10; Length 563;  
 Best Local Similarity 41.6%; Pred. No. 1.8e-24;  
 Matches 79; Conservative 40; Mismatches 56; Indels 15; Gaps 4;

QY 7 YPIINFTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRGLPINQRFILVELSNH 66  
 DB 28 YPSVSNLAGAKSATYRDFLKNLRITVATGYEVNGLPVLRRSEVQVKRNFVLVLLTNY 87

QY 67 AELSVTALDVTNAYVYVGRAGNSAYFFHPDQDEAEI--THLFTDVQNRVYFAFGNY 124  
 DB 88 NGNTVTILAVDTNLYVAFSANANSYFF-----KDATQLOKSNLFGVTR-QHTLPTFTGNY 141

QY 125 DRLEQAGNLRNIELNGPLGPLEEAISSALYYSTGGTQLTPLARSFIIQMISEARFOY 184  
 DB 142 DNLTAAGTRESIELGSPIDGNAITSYDE-----SVARSLVVIQWSEARFRY 194

QY 185 IEGEMTRIR 194  
 DB 195 IEQEVRSLSQ 204

RESULT 7  
 Q9AVR2 PRELIMINARY; PRT; 564 AA.  
 AC Q9AVR2  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein precursor (EC 3.2.2.22) (rRNA N-glycosidase).  
 GN EBU1.  
 OS Sambucus ebulus.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=28503;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Girbes T., Iglesias R., Perez Y., Ferreras J.M., Citores L.;  
 RT "Molecular cloning of ebulin I.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AJ400822; CAC33178.1; --  
 DR HSSP; P02879; 2AAL.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00438; RICIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Glycosidase; Hydrolase; Signal; Toxin.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 298 EBU1IN L A-CHAIN.  
 FT CHAIN 299 564 EBU1IN L B-CHAIN.  
 SQ SEQUENCE 564 AA; 62694 MW; 8261681A6DB55CB8 CRC64;

Query Match 33.9%; Score 347.5; DB 10; Length 564;  
 Best Local Similarity 41.1%; Pred. No. 3.4e-24;  
 Matches 78; Conservative 41; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRGLPINQRFILVELSNH 66  
 DB 28 YPSVSNLAGAKSATYRDFLKNLRITVATGYEVNGLPVLRRSEVQVKRNFVLVLLTNY 87

QY 67 AELSVTALDVTNAYVYVGRAGNSAYFFHPDQDEAEI--THLFTDVQNRVYFAFGNY 124  
 DB 88 NGDTVTSADVDTNLYVAFSANGNSYFF-----KQATELOKSNLFLGT-QHTLSFTGNY 141

Query Match 33.6%; Score 344.5; DB 10; Length 528;  
 Best Local Similarity 45.0%; Pred. No. 5.9e-24;  
 Matches 85; Conservative 24; Mismatches 71; Indels 9; Gaps 4;

QY 6 QYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHPIVLPNVRGLPINQRFILVELSN 65  
 DB 1 QDQVIKETTEGATSQSYKQFIEALRQLTGG--LIHDIPLVLPDPTTVEERNRYITVELSN 58

QY 66 HAEISVTALDVTNAYVYVGRAGNSAYFFHPDQDEAEI--THLFTDVQNRVYFAFGNY 125  
 DB 59 SERESIEVGIDVTNAYVYVGRAGNSQSYFL---RDAPASATYLPFGTQ-RYSLRFDGSYG 114

QY 126 RLEQAGNLRNIELNGPLGPLEEAISSALYYSTGGTQLTPLARSFIIQMISEARFOYI 185  
 DB 115 DLEWHAQTRIEISLGLQALHAIS---FLRSGANDSEKARTLIVIIQASEARFRCI 171

QY 186 EGEMLTRIR 194  
 DB 172 SNRVGVSR 180

RESULT 9  
 Q94KE4 PRELIMINARY; PRT; 289 AA.  
 ID Q94KE4  
 AC Q94KE4;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Trichosanthen precursor (EC 3.2.2.22) (rRNA N-glycosidase).  
GN TCS.  
OS Trichosanthos kirilowii (Mongolian snake-gourd).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
OX NCBI\_TaxID=3677;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yuan H., Wang L., Wang Y., An C., Chen Z.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC -!- SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AF367252; AAK52960.1; -  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Signal; Toxin.  
FT SIGNAL 1 23  
FT CHAIN 24 270 TRICHOSANTHIN.  
FT CHAIN 24 270  
SQ SEQUENCE 289 AA; 31706 MW; A6D5602549CA5657 CRC64;  
  
Query Match 33.2%; Score 340; DB 10; Length 289;  
Best Local Similarity 39.5%; Pred. No. 6.8e-24;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
  
QY 10 INFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVLSNHAEL 69  
DB 25 VSFRLSGATSSSYGVFISNRKALPNERKL-YDIPLL--RSSLPQSQRVALHILNTYADE 81  
QY 70 SVTLALDVTNAYVVGVRAGNSAYFFHPDQEDA-EAITHLFTDVQNRYYTFAFGNYDRLE 128  
DB 82 TISVAIDVTNVIMGYRAGDTSYFF--NEASATEAAKVVKDKMRKVTLPYSGNYERLQ 138  
QY 129 QLAGNLRNIELNGPLBESALYYSTGGTQPLTARSFFICIMISEARFQVIEGE 188  
DB 139 TAAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMVLIQSTSEAAARYKFEIQ 193  
QY 189 MRTRI 193  
DB 194 IGRV 198  
  
RESULT 10  
Q9LRE3  
ID Q9LRE3 PRELIMINARY; PRT; 247 AA.  
AC Q9LRE3  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).  
GN TBK.  
OS Trichosanthos sp. Bac Kan 8-98.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
OX NCBI\_TaxID=118182;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;  
RT "Genomic DNA clone for mature typ-1 ribosome-inactivating protein from  
RT Trichosanthos sp. sample 01 Bac Kan 8-98 Vien CNH (Hanoi).";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC -!- SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AB039324; BAA92530.1; -  
DR HSSP; P09989; 1MRJ.  
DR InterPro; IPR001574; RIP.

DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Toxin.  
FT NON\_TER 1 247  
FT NON\_TER 247 247  
SQ SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;  
  
Query Match 33.0%; Score 338; DB 10; Length 247;  
Best Local Similarity 40.0%; Pred. No. 8.5e-24;  
Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5;  
  
QY 10 INFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVLSNHAEL 69  
DB 2 VSFRLSGATSSSYGVFISNRKALPNERKL-YDIPLL--RSTLPQSQRVALHILNTYADE 58  
QY 70 SVTLALDVTNAYVVGVRAGNSAYFFHPDQEDA-EAITHLFTDVQNRYYTFAFGNYDRLE 128  
DB 59 TISVAIDVTNVIMGYRAGDTSYFF--NEASATEAAKVVKDKMRKVTLPYSGNYERLQ 115  
QY 129 QLAGNLRNIELNGPLBESALYYSTGGTQPLTARSFFICIMISEARFQVIEGE 188  
DB 116 IAAGKIRENIPGLPALDSAITTLFYNNAN-----SAASALMVLIQSTSEAAARYKFEIQ 170  
QY 189 MRTRI 193  
DB 171 IGRV 175  
  
RESULT 11  
Q41216  
ID Q41216 PRELIMINARY; PRT; 289 AA.  
AC Q41216  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).  
GN TRICHOSANTHIN\_TCS  
OS Trichosanthos kirilowii (Mongolian snake-gourd).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
OX NCBI\_TaxID=3677;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94271613; PubMed=8003348;  
RA Zheng H., Wang B., Shaw P., Yeung H.;  
RT "[Cloning and DNA sequencing of the gene encoding trichosanthin].";  
RL 1 Chuan Huah Pao 21:42-51(1994).  
CC -!- CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC -!- SPECIFIC ADENOSINE ON THE 28S RRNA.  
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; S70176; AAB31048.1; -  
DR HSSP; P09989; 1MRJ.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Hydrolase; Toxin.  
SQ SEQUENCE 289 AA; 31650 MW; 286AC14D48BCA175 CRC64;  
  
Query Match 33.0%; Score 338; DB 10; Length 289;  
Best Local Similarity 39.5%; Pred. No. 1.1e-23;  
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;  
  
QY 10 INFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVLSNHAEL 69  
DB 25 VSFRLSGATSSSYGVFISNRKALPNERKL-YDIPLL--RSSLPQSQRVALHILNTYADE 81  
QY 70 SVTLALDVTNAYVVGVRAGNSAYFFHPDQEDA-EAITHLFTDVQNRYYTFAFGNYDRLE 128  
DB 82 TISVAIDVTNVIMGYRAGDTSYFF--NEASATEAAKVVKDKMRKVTLPYSGNYERLQ 138

QY 129 OLAGNRENIELGNGPLEEALSYLYSTGGTQGLPTLARSFIICIMISEARFQIEG 188  
 Db 139 TAAGKIRENIPGLPALDSAITLIFYNNAN-----SAASALMVLIOSTSEARFKIEQ 193  
 QY 189 MRTRI 193  
 Db 194 IGKRV 198

RESULT 12  
 Q38760 ID Q38760 PRELIMINARY; PRT; 252 AA.  
 AC Q38760;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Abirin-E (RRNA N-glycosidase) (EC 3.2.2.22) (Fragment).  
 GN RIP.  
 OS Abus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OC NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RX MEDLINE=91201329; PubMed=2016300;  
 RA Evensen G., Mathiesen A., Sundan A.;  
 RT "direct molecular cloning and expression of two distinct abrin A-  
 chains.";  
 RL J. Biol. Chem. 266:6848-6852(1991).  
 CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
 FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
 PRECEDES ENDOCYTOSIS.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -!- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING  
 PROTEINS. BELONGS TO TYPE 2 RIP.  
 CC EMBL; X54872; CRA38654.1; -;  
 DR HSPF; P11140; IABR.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.  
 FT CHAIN 1 252 ABRIN E, A CHAIN (BY SIMILARITY).  
 FT NON\_TER 252 252  
 SQ SEQUENCE 252 AA; 28309 MW; BBFC846B9E92B5DE CRC64;

Query Match 32.8%; Score 336.5; DB 10; Length 252;  
 Best Local Similarity 44.9%; Pred. No. 1.2e-23;  
 Matches 83; Conservative 24; Mismatches 69; Indels 9; Gaps 4;

QY 10 INFRTAGATVQSYNTFIRAVRGRLTTCADVRHEIPVLPNRVGLPINORFILVELSNHAE 69  
 Db 6 IKFSTEGATSQSYKQFIEALRERLGG--LIHDIPLVRDPTTVBERNRYITVELSNSE 63  
 QY 70 SVTLALDVNAYVGYRAGNSAYFFHPDNOEDAEAI--THLFTDVQNRYYTFAFGNVDLEQ 129  
 Db 64 SIEVGIDVNAVYRAGNSQSYFL---RDPASASTYLTGTQ-RYSLRFDGSGYGLER 119  
 QY 130 LAGNRENIELGNGPLEEALSYLYSTGGTQGLPTLARSFIICIMISEARFQIEG 189  
 Db 120 WAHQTRQISLGLQALTHAIS---FLRSGANDEKARTLIVIOQASEARRYISNRV 176  
 QY 190 RTRIR 194

Db 177 GVSIR 181

RESULT 13  
 Q945S2 ID Q945S2 PRELIMINARY; PRT; 563 AA.  
 AC Q945S2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
 N-glycosidase).  
 GN AVL.  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OC NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RA Van Damme E.J.M.;  
 RT "Characterization and cloning of lectins and ribosome-inactivating  
 proteins from Sambucus nigra leaves.";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 SPECIFIC ADENOSINE ON THE 28S RRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AF409135; AAL04123.1; -;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; Ricin.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Hydrolase; Toxin.  
 SQ SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;

Query Match 32.2%; Score 330.5; DB 10; Length 563;  
 Best Local Similarity 39.5%; Pred. No. 1.4e-22;  
 Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;

QY 7 YPIINFRTAGATVQSYNTFIRAVRGRLTTCADVRHEIPVLPNRVGLPINORFILVELSNH 66  
 Db 28 YPSVSFNLDCAKSATYRDFLSNLRKTVATGTYEVNGLPVLRRSEVOVKSRFVLPLTNY 87  
 QY 67 AELSVTALDVNAYVGYRAGNSAYFFHPDNOEDAEAI--THLFTDVQNRYYTFAFGN 124  
 Db 88 NGNTVTILADVNTLNYVAFSGNANSYFF-----KDTEVQKSNLFVGTQKN-TLSFTGNY 141  
 QY 125 DRLEQLAGNRENIELGNGPLEEALSYLYSTGGTQGLPTLARSFIICIMISEARFQY 184  
 Db 142 DNLETAANTRESEIELGSPDLGAI--SVARSLLVVIQMVSEARFRY 194  
 QY 185 IEGEMRTRIR 194  
 Db 195 IEGEVRSLQ 204

RESULT 14  
 Q8GT32 ID Q8GT32 PRELIMINARY; PRT; 563 AA.  
 AC Q8GT32;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein nigrin 1 precursor  
 (EC 3.2.2.22).  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:20 ; Search time 32.2239 Seconds  
(without alignments)  
930.966 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MIFPKQVPIINFTAGATVQ.....ARFQYIEGEMTRVYNRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:\*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	951	97.9	267	14	AA1980.DAT:*
2	951	97.9	267	16	AA1981.DAT:*
3	951	97.9	290	18	AA1982.DAT:*
4	951	97.9	290	18	AA1983.DAT:*
5	951	97.9	332	8	AA1984.DAT:*
6	951	97.9	332	8	AA1985.DAT:*
7	951	97.9	332	10	AA1986.DAT:*
8	951	97.9	554	16	AA1987.DAT:*
9	951	97.9	562	10	AA1988.DAT:*

10	951	97.9	565	6	AA1989.DAT:*
11	951	97.9	565	22	AA1990.DAT:*
12	951	97.9	565	22	AA1991.DAT:*
13	951	97.9	576	8	AA1992.DAT:*
14	951	97.9	576	18	AA1993.DAT:*
15	951	97.9	576	20	AA1994.DAT:*
16	951	97.9	576	21	AA1995.DAT:*
17	951	97.9	576	22	AA1996.DAT:*
18	951	97.9	576	22	AA1997.DAT:*
19	949	97.7	268	14	AA1998.DAT:*
20	947	97.5	200	9	AA1999.DAT:*
21	946	97.4	534	14	AA2000.DAT:*
22	946	97.4	565	7	AA2001.DAT:*
23	944	97.2	574	8	AA2002.DAT:*
24	942	97.0	574	10	AA2003.DAT:*
25	941	96.9	267	13	AA2004.DAT:*
26	941	96.9	267	21	AA2005.DAT:*
27	940	96.8	332	11	AA2006.DAT:*
28	934	96.2	267	16	AA2007.DAT:*
29	918	94.5	267	14	AA2008.DAT:*
30	861.5	88.7	540	18	AA2009.DAT:*
31	861.5	88.7	540	18	AA2010.DAT:*
32	801.5	82.5	534	8	AA2011.DAT:*
33	718	73.9	280	10	AA2012.DAT:*
34	336	34.6	247	16	AA2013.DAT:*
35	336	34.6	247	21	AA2014.DAT:*
36	336	34.6	248	11	AA2015.DAT:*
37	336	34.6	248	13	AA2016.DAT:*
38	336	34.6	267	18	AA2017.DAT:*
39	336	34.6	267	18	AA2018.DAT:*
40	336	34.6	289	11	AA2019.DAT:*
41	336	34.6	289	13	AA2020.DAT:*
42	336	34.6	289	13	AA2021.DAT:*
43	336	34.6	289	14	AA2022.DAT:*
44	336	34.6	289	15	AA2023.DAT:*
45	336	34.6	289	18	AA2024.DAT:*

#### ALIGNMENTS

##### RESULT 1

AA1980.DAT:\*

ID AA1980.DAT:\*

AC AA1980.DAT:\*

XX AA1980.DAT:\*

XX AA1980.DAT:\*

DT 25-MAR-2003 (updated)

DT 09-JAN-2003 (updated)

DT 13-SEP-1993 (first entry)

DE Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

XX Ricin A chain.

Sequence of prepro  
Castor bean prepro  
Modified castor be  
Sequence of Ricinu  
Castor bean ricin  
Castor bean ricin  
Ricin A chain (RTA  
Ricin A chain (RTA  
Ricin A. Syntheti  
Castor oil plant a  
R. communis agglut  
Sequence of Ricin  
Ricin agglutinin A  
Trichosanthin anti  
Amino acid sequenc  
Synthetic alpha-tr  
Mature alpha-Trich  
Trichosanthin (a r  
Trichosanthin. Tr  
Trichosanthin from  
Trichosanthin prot  
Trichosanthin cu  
Encodes chinese cu  
Alpha-trichosanthi  
Chinese cucumber a

Type II ribosome-inactivating protein; type II RIP; gelonin; toxin.  
momordin; immunoconjugate; autoimmune disease; cell killing; toxin.

Unidentified.  
WO9309130-A1.  
13-MAY-1993.  
04-NOV-1992; 92WO-US09487.  
04-NOV-1991; 91US-0787567.  
19-JUN-1992; 92US-0901707.  
(XOMA ) XOMA CORP.  
Berhard SL, Better MD, Carroll SF, Lane JA, Lei SP;  
WPI; 1993-167617/20.

PT Analogues of type I ribosome inactivating protein - useful as  
PT cytotoxic agents, immuno toxins for treating autoimmune diseases,  
XX cancer, graft versus host disease and selective cell killing in-vivo  
XX  
XX Claim 1; Page 92; 163pp; English.  
XX  
CC The invention covers analogues of Type I RIPS. Ricin is a Type II  
CC RIP whose A chain is homologous to plant type I RIPS. The analogues  
CC of the invention have a cysteine available for intermolecular  
CC disulphide bonding at an amino acid position corresp. to a position  
CC not naturally available for bonding; the cys residue is located in  
CC the C-terminal region of the analogue between a position corresp. to  
CC amino acid 251 and the C-terminus of ricin A chain. The analogues are  
CC pref. joined via a disulphide linkage to a molecule which specifically  
CC binds to a target cell, e.g. an antibody fragment.  
CC (Updated on 09-JAN-2003 to add missing OS field.)  
XX  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 267 AA;

Query Match 97.9%; Score 951; DB 14; Length 267;  
Best Local Similarity 94.9%; Pred. No. 1.3e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 111  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 171  
DB 121 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 180  
QY 172 FOYIEGEMRTRIRYNRS 189  
DB 181 FOYIEGEMRTRIRYNRS 198

RESULT 2  
AAR63902  
ID AAR63902 standard; protein; 267 AA.  
XX  
AC AAR63902;  
XX  
XX 25-MAR-2003 (updated)  
DT 27-JUL-1995 (first entry)  
XX  
XX Ricin A-chain (RTA).  
XX  
XX Ricin A chain; RTA; ribosome-inactivating proteins; RIPS;  
XX cytotoxic therapeutic agents; autoimmune disease; cancer;  
XX graft-versus-host disease.  
XX  
XX Ricinus communis.  
XX  
XX WO9426910-A1.  
XX  
XX 24-NOV-1994.  
XX  
XX 12-MAY-1994; 94WO-US05348.  
XX  
XX 12-MAY-1993; 93US-0064691.  
XX  
XX (XOMA ) XOMA CORP.  
XX  
XX Better MD, Carroll SS, Studnicka GM, Carroll SF;  
XX  
XX WPI; 1995-006804/01.  
XX  
XX Polynucleotide(s) encoding type I ribosome-inactivating proteins

PT - which are suitable for use as components of cytotoxic  
PT therapeutic agents.  
XX  
XX Example 3; Fig 1; 221pp; English.  
XX  
CC AAR63902 is the ricin A chain gene product, it is analogous to the  
CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.  
CC RIPS are the key components of cytotoxic therapeutic agents (CTAs),  
CC which include gene fusion products and immunoconjugates. CTAs may  
CC be used to selectively eliminate any cell type to which a RIP  
CC component is targeted, by the specific binding capacity of the  
CC second component of the agent. They can be used in the treatment  
CC of diseases where the elimination of a particular cell type is  
CC desired, such as autoimmune disease, cancer and graft-versus-host  
CC disease.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 267 AA;

Query Match 97.9%; Score 951; DB 16; Length 267;  
Best Local Similarity 94.9%; Pred. No. 1.3e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 111  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 171  
DB 121 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 180  
QY 172 FOYIEGEMRTRIRYNRS 189  
DB 181 FOYIEGEMRTRIRYNRS 198

RESULT 3  
AAR63902  
ID AAR63902 standard; protein; 290 AA.  
XX  
AC AAR63902;  
XX  
XX 25-MAR-2003 (updated)  
DT 02-DEC-1997 (first entry)  
XX  
XX Ricin A-chain ribosome inhibitory protein inactive precursor.  
XX  
XX Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;  
XX internal linker; Barley translation inhibitor; trichosanthin;  
XX Ricin A-chain; Abrin-A chain; Saporin; SUT-1; Luffin A; MAP;  
XX Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;  
XX therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;  
XX post-translational modification; cancer; neoplasia; HIV; AIDS;  
XX human immunodeficiency virus; acquired immune deficiency syndrome.  
XX  
XX Synthetic.  
XX  
XX US5646026-A.  
XX  
XX 08-JUL-1997.  
XX  
XX 07-JUN-1995; 95US-0485286.  
XX  
XX 09-DEC-1992; 92US-0987927.  
XX  
XX 11-JUN-1990; 90US-0535636.  
XX  
XX 26-JAN-1995; 95US-0378761.  
XX  
XX 07-JUN-1995; 95US-0485286.  
XX  
XX (DOWC ) DOWELANCO.

XX Hey TD, Morgan AER, Walsh TA;  
 XX WPI; 1997-362934/33.  
 XX DNA encoding pro-ribosome inactivating proteins - inactive  
 XX precursors of ribosome inactivating proteins; can be expressed in  
 XX eukaryotic cells without causing cell death  
 XX Claim 4; Column 91-94; 186pp; English.  
 XX AAW25136 represents a Ricin A-chain ribosome inhibitory protein (RIP)  
 XX which was engineered to contain a selectively removable internal peptide  
 XX linker sequence separating the alpha and beta units of the RIP. When  
 XX separated the two units regain activity and are capable of inactivating  
 XX eukaryotic ribosomes and hence preventing protein production. Many  
 XX different RIPs may be produced with an internal linker including  
 XX maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and  
 XX Saporin. The RIPs can be used in the construction of therapeutic  
 XX toxins targeted to specific cells such as tumour cells via the  
 XX attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
 XX A further use is in HIV therapy (see US4869903). There is interest  
 XX in expressing RIP recombinantly in host eukaryotic cells, because of  
 XX the capacity to provide correct post-translational processing. However,  
 XX RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
 XX in cell death. Since the inactive RIP proteins are not cytotoxic to  
 XX eukaryotic cells, they can be recombinantly expressed in such cells and  
 XX then converted to active RIP proteins.  
 XX (Updated on 25-MAR-2003 to correct PF field.)  
 XX Sequence 290 AA;  
 XX Query Match 97.9%; Score 951; DB 18; Length 290;  
 XX Best Local Similarity 94.9%; Pred. No. 1.4e-93;  
 XX Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLIV 51  
 Db |||||  
 QY 25 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLTGGADVVRHEIPVLPNRVGLPINORFLIV 84  
 Db |||||  
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFG 111  
 Db |||||  
 QY 85 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFG 144  
 Db |||||  
 QY 172 FQYIEGEMRTRIRYNRRS 189  
 Db |||||  
 QY 205 FQYIEGEMRTRIRYNRRS 222  
 Db |||||  
 RESULT 4  
 AAW21699  
 ID AAW21699 standard; Protein; 290 AA.  
 XX AAW21699;  
 XX 25-MAR-2003 (updated)  
 XX 26-SEP-1997 (first entry)  
 XX Ricin A-chain RIP.  
 XX pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;  
 XX inactivation; eukaryotic ribosome; alpha fragment; beta fragment;  
 XX inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;  
 XX rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.  
 XX Ricinus communis.  
 XX Key Location/Qualifiers  
 XX FH 152..162  
 XX Region

FT /note= "Position of possible insertion of internal  
 FT peptide linker sequence"  
 XX US5635384-A.  
 XX 03-JUN-1997.  
 XX 26-JAN-1995; 95US-0378761.  
 XX 09-DEC-1992; 92US-0987927.  
 XX 11-JUN-1990; 90US-0535636.  
 XX 26-JAN-1995; 95US-0378761.  
 XX (DOWC ) DOWELANCO.  
 XX Hey TD, Morgan AER, Walsh TA;  
 XX WPI; 1997-309831/28.  
 XX Inactive precursor of maize ribosome-inactivating protein - also  
 XX chimeric ribosome-inactivating protein precursors containing  
 XX internal linker sequences  
 XX Claim 2; Column 91-94; 121pp; English.  
 XX The sequences given in AAW21698-710 represent Ribosome Inactivating  
 XX Proteins (RIP's), which may be used in the construction of the  
 XX proRIP of the invention. The proRIP has a selectively removable,  
 XX internal peptide linker. The precursor sequence is incapable of  
 XX inactivating eukaryotic ribosomes, but can be converted by removal  
 XX of the linker into a protein having alpha and beta fragments and being  
 XX capable of inactivating eukaryotic ribosomes. RIPs are potent  
 XX inhibitors of eukaryotic protein synthesis. They possess a highly  
 XX specific N-glycosidase activity which cleaves the glycosidic bond of  
 XX adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit  
 XX cellular proliferation of cells, e.g. cancer cells and HIV-infected T  
 XX cells. The inactive proRIP proteins make it possible to provide protein  
 XX synthesis inhibitors with uses in practical and improved ways not before  
 XX possible. The RIP can be used to make cytotoxic conjugates.  
 XX (Updated on 25-MAR-2003 to correct PF field.)  
 XX Sequence 290 AA;  
 XX Query Match 97.9%; Score 951; DB 18; Length 290;  
 XX Best Local Similarity 94.9%; Pred. No. 1.4e-93;  
 XX Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLIV 51  
 Db |||||  
 QY 25 IFPKQYPIINFNTAGATVQSYTNFIRAVRGRLTGGADVVRHEIPVLPNRVGLPINORFLIV 84  
 Db |||||  
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFG 111  
 Db |||||  
 QY 85 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFG 144  
 Db |||||  
 QY 112 GNYDRLEQLAGNLRENIELGNGLPEEAISALYYSTGTQPTLARSFIIQMISEAR 171  
 Db |||||  
 QY 145 GNYDRLEQLAGNLRENIELGNGLPEEAISALYYSTGTQPTLARSFIIQMISEAR 204  
 Db |||||  
 QY 172 FQYIEGEMRTRIRYNRRS 189  
 Db |||||  
 QY 205 FQYIEGEMRTRIRYNRRS 222  
 Db |||||  
 RESULT 5  
 AAP70097  
 ID AAP70097 standard; protein; 332 AA.  
 XX AAP70097;  
 XX 09-APR-1991 (first entry)  
 XX Ricin A.



DT 13-AUG-1990 (first entry)  
XX Ricin A encoded by insert from plasmid pRA123.  
DE Plasmid pRA123; ricin-A; ricin-B; cytotoxicity.  
XX Ricinus communis.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Peptide 1...35  
FT Peptide /label= leader sequence  
FT Peptide 36..302  
FT Peptide /label=A-chain  
FT Peptide 303..314  
FT Peptide /label=linker  
FT Peptide 315..332  
FT Peptide /label=B-chain  
PN EP335476-A.  
XX  
XX 04-OCT-1989.  
PD  
XX 19-JAN-1989; 89EP-0201162.  
XX  
XX 08-FEB-1984; 84US-0578115.  
PR 08-FEB-1984; 84US-0578121.  
PR 08-FEB-1984; 84US-0578122.  
PR 07-SEP-1984; 84US-0648759.  
PR 20-SEP-1984; 84US-0653515.  
XX  
XX (CETU ) CETUS CORPORATION.  
PA  
XX Gelfand D, Lawyer PC, Horn G, Greenfield L, Nitecki D, Kaplan D;  
PI Piatak MJ;  
XX  
XX WPI; 1989-286959/40.  
DR N-PSDB; AAN91281.  
XX  
XX Recombinant vectors expressing ricin chains or diphtheria toxin -used for  
PT prodn. of new immunotoxin conjugates with monoclonal antibodies, having  
PT high cell specificity and good extracellular stability.  
XX  
XX Disclosure; Fig 14; 54pp; English.  
XX  
XX Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for  
CC ricin A, as well as codons for 12 AAs joining the A to the B chain.  
CC Following modification for ease of manipulation the plasmid was used to  
CC construct expression vectors which express the conjugates in  
CC host cells.  
CC (Updated on 31-OCT-2002 to add missing OS field.)  
CC (Updated on 25-MAR-2003 to correct PF field.)  
CC (Updated on 25-MAR-2003 to correct PI field.)  
XX  
XX Sequence 332 AA;  
SQ  
Query Match 97.9%; Score 951; DB 10; Length 332;  
Best Local Similarity 94.9%; Pred. No. 1.7e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFITAGATVQSYNTFIRAVRGRLT-----VLPNRVGLPQNFILV 51  
DB 36 IFPKQYPIINFITAGATVQSYNTFIRAVRGRLTGGADVRHEIPVLPNRVGLPQNFILV 95  
QY 52 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFPHDPNQEDAEAITHLFTDVQNRVYTFAG 111  
DB 96 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFPHDPNQEDAEAITHLFTDVQNRVYTFAG 155  
QY 112 GNYDRLEQAGNLRNENIELNGNPLEEALISALYYSTGGTGLPTLARSFFICIQMISEAAR 171  
DB 156 GNYDRLEQAGNLRNENIELNGNPLEEALISALYYSTGGTGLPTLARSFFICIQMISEAAR 215

QY 172 FOYIEGEMTRIRYNRRS 189  
DB 216 FOYIEGEMTRIRYNRRS 233  
RESULT 8  
AAR70827  
ID AAR70827 standard; Protein; 554 AA.  
XX  
XX AAR70827;  
AC  
XX 25-MAR-2003 (updated)  
DT 31-AUG-1995 (first entry)  
XX  
XX Anti-cataract immunotoxin.  
XX Immunotoxin; heavy chain; light chain; variable region; antibody;  
KW ricin-A; cytostatic; cataract; lens opacification; epithelial cell;  
KW pHB19; 4197X; monoclonal antibody; Mab.  
XX  
XX Synthetic.  
XX  
XX Location/Qualifiers  
FH Key 1...27  
FT Peptide /label= Sig\_peptide  
FT /note= "phoA signal sequence"  
FT Domain 28..145  
FT /label= HEAVY  
FT /note= "Mab 4197X heavy chain"  
FT Peptide 148..166  
FT /label= LINKER  
FT Domain 169..274  
FT /label= LIGHT  
FT /note= "Mab 419X light chain"  
FT Domain 276..544  
FT /label= RICIN-A  
FT Peptide 549..554  
FT /label= TAG  
FT /note= "hexa-histidine tail"  
XX  
XX WO9503828-A1.  
XX  
XX 09-FEB-1995.  
XX  
XX 15-JUL-1994; 94WO-US07919.  
XX  
XX 02-AUG-1993; 93US-0101329.  
XX  
XX (HOUS-) HOUSTON BIOTECHNOLOGY INC.  
XX  
XX Gould RM, Kelleher PJ, Wallace TL, Wood MS;  
PI  
XX WPI; 1995-082036/11.  
DR N-PSDB; AAQ85386.  
XX  
XX New single chain immunotoxin - binds specifically to epithelial  
PT cells, for inhibiting development of sec. cataracts after  
PT extra: capsular cataract extraction.  
XX  
XX Disclosure; Fig.4; 68pp; English.  
XX  
XX The immunotoxin given in AAR70827 comprises the heavy and light chain  
CC variable regions of anti-lens epithelium IgG3 Mab 4197X linked to  
CC ricin-A and a hexa-histidine tag. The DNA construct encoding the  
CC immunotoxin was expressed from pHB19 in E. coli.  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 554 AA;  
SQ  
Query Match 97.9%; Score 951; DB 16; Length 554;  
Best Local Similarity 94.9%; Pred. No. 3.4e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 278 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 337  
 QY 52 ELSNHAELS VTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 111  
 DB 338 ELSNHAELS VTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 397  
 QY 112 GNYDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQLP TARSFFIICQMISEAAR 171  
 DB 398 GNYDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQLP TARSFFIICQMISEAAR 457  
 QY 172 FOYIEGEMTRIRYNRRS 189  
 DB 458 FOYIEGEMTRIRYNRRS 475  
 RESULT 9  
 AAP90079 standard; protein; 562 AA.  
 XX  
 AC AAP90079;  
 DT 25-MAR-2003 (updated)  
 DT 01-NOV-1989 (first entry)  
 DE Ricin D.  
 XX Ricin D; Ricinus communis; castor beans; Zanicbariensis variety;  
 KW modified; lectin binding removed; reduced cell binding  
 XX Ricinus communis (castor beans).  
 OS  
 PN WO8904839-A.  
 XX  
 PD 01-JUN-1989.  
 XX  
 PF 23-NOV-1988; 88WO-US04238.  
 XX  
 PR 24-NOV-1987; 87US-0124735.  
 XX  
 PA (GEMY) GENETICS INST INC.  
 XX  
 PI Brown EL, Jones S;  
 XX  
 DR WPI; 1989-178366/24.  
 DR N-PSDB; AAN90068.  
 XX  
 PT Modified ricin molecules and toxin conjugates  
 PT in which the lectin binding function of the B chain  
 PT is removed or diminished to reduce cell binding.  
 XX  
 PS Disclosure; fig 1; 51pp; English.  
 XX  
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment  
 CC of DNA from Ricinus communis, Zanicbariensis variety. Patent  
 CC discloses many modifications of ricin in which the lectin binding  
 CC function of the B chain is diminished or removed, and conjugation  
 CC to toxins to eliminate cell binding.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 562 AA;  
 Query Match 97.9%; Score 951; DB 10; Length 562;  
 Best Local Similarity 94.9%; Pred. No. 3.5e-93;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
 DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELS VTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 111

DB 96 ELSNHAELS VTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQLP TARSFFIICQMISEAAR 171  
 DB 156 GNYDRLEQLAGNLRNIELNGPLLEEAISALYYSTGGTQLP TARSFFIICQMISEAAR 215  
 QY 172 FOYIEGEMTRIRYNRRS 189  
 DB 216 FOYIEGEMTRIRYNRRS 233  
 RESULT 10  
 AAP50166 standard; Protein; 565 AA.  
 XX  
 AC AAP50166;  
 DT 16-OCT-1991 (first entry)  
 XX  
 DE Sequence of preprorin encoded by pRCL617.  
 XX  
 KW Toxin; anti-tumour therapy.  
 XX  
 OS Ricinus.  
 XX  
 PH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /label= signal  
 FT Protein 25..565  
 FT Region 292..303  
 FT /label= links the C-terminus of the A chain and  
 FT the N-terminus of the B chain  
 FT Modified-site 34..36  
 FT /label= N-linked glycosylation  
 FT Modified-site 260..262  
 FT /label= N-linked glycosylation  
 FT Modified-site 398..400  
 FT /label= N-linked glycosylation  
 FT Modified-site 438..440  
 FT /label= N-linked glycosylation  
 XX  
 PN EPI45111-A.  
 XX  
 PD 19-JUN-1985.  
 XX  
 PF 13-JUL-1984; 84EP-0304801.  
 XX  
 PR 13-MAR-1984; 84GB-0006569.  
 PR 15-JUL-1983; 83GB-0019265.  
 PR 15-JUL-1983; 83CH-0019265.  
 XX  
 PA (UFWA-) UNIV WARWICK.  
 XX  
 PI Lord JM, Roberts LM, Lamb FI;  
 XX  
 DR WPI; 1985-148040/25.  
 DR N-PSDB; AAN50202.  
 XX  
 PT New DNA sequences coding for ricin type plant toxin - or its  
 PT mutants, and modified vectors and host microorganisms  
 XX  
 PS Disclosure; Page 30-30c; 40pp; English.  
 XX  
 CC Preprorin is the whole polypeptide encoded by AAN50202 and the DNA  
 CC encoding this is claimed. Preprorin is obtained from preprorin by  
 CC removal of the AA leader sequence. The linker AA sequence which is  
 CC present in the precursor polypeptide is enzymatically removed in the  
 CC cell to separate the A and B chains, which are joined by a  
 CC disulphide bridge during the formation of the ricin molecule itself.  
 CC This linker region as well as the presumptive amino terminal leader  
 CC or signal sequence are not present in the sequences already  
 CC published by Funatsu et al.

XX Sequence 565 AA;  
Query Match 97.9%; Score 951; DB 6; Length 565;  
Best Local Similarity 94.9%; Pred. No. 3.5e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IPKQYPIINFITAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 51  
DB 25 IPKQYPIINFITAGATVQSYNFIKAVRGLTGGADVRHDPVLPNRVGLPINQRFILV 84  
QY 52 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAPG 111  
DB 85 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAPG 144  
QY 112 GNYDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 171  
DB 145 GNYDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 204  
QY 172 FOYIEGEMTRIRYNRRS 189  
DB 205 FOYIEGEMTRIRYNRRS 222  
RESULT 11  
AAG78300  
ID AAG78300 standard; Protein; 565 AA.  
XX AAG78300;  
XX 15-NOV-2001 (first entry)  
XX Castor bean preproricin protein (SEQ ID 1).  
XX Castor bean plant; preproricin; ricin; A chain; B chain;  
KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
KW retroviral infection; anti-HIV; virucide activity; viral protease.  
XX Ricinus communis.  
XX Key Location/Qualifiers  
FH Peptide 1..24  
FT /label= Signal peptide  
FT Protein 25..290  
FT /label= Ricin A chain  
FT /note= "N-glycosidase"  
FT Peptide 291..302  
FT /label= Linker\_peptide  
FT /note= "Cleaved during activation of ricin"  
FT Protein 303..565  
FT /label= Ricin B chain  
FT /note= "Galactose/N-acetylgalactosamine-binding lectin"  
XX WO200160393-A1.  
XX 23-AUG-2001.  
XX 15-FEB-2001; 2001WO-US05282.  
XX 16-FEB-2000; 2000US-0182759.  
XX (BECH-) BECHTEL BWXT IDAHO LLC.  
XX Keener WK, Ward TE;  
XX WPI; 2001-581908/65.  
XX N-PSDB; AAI64137.  
XX Novel composition comprising toxin e.g., ricin based antiviral compound  
PT useful for treating viral infections such as human immunodeficiency  
PT virus infection.  
XX Disclosure; Page 47-50; 66pp; English.

XX The sequence relates to preproricin protein encoded by the DNA sequence  
CC given in AAI64137. The invention relates to a novel toxin (e.g., ricin)  
CC based antiviral agent which is toxic to virus-infected cells, but  
CC non-toxic to uninfected cells. The invention has anti-HIV and virucide  
CC activities. Its mechanism of action is through inactivation of cellular  
CC ribosomes and enhancement of binding of the antiviral agent to galactose  
CC residues on cell surfaces, and its cellular internalisation. The  
CC invention is useful for treating human immunodeficiency virus infection  
CC and other viral infections, especially retroviral infections. The  
CC antiviral agent is activated in viral particles or early-stage infected  
CC cells, killing the cells upon infection and effectively preventing the  
CC integration of the viral genome into the host genome thereby preventing  
CC the latency/rebound problem. The agent enters all HIV susceptible cells,  
CC and not just cells known to act as host cells for the virus. The  
CC antiviral agent remains inert in a cell until degraded in it, unless the  
CC cell is infected with the virus, where the viral protease activates it.  
XX Sequence 565 AA;  
Query Match 97.9%; Score 951; DB 22; Length 565;  
Best Local Similarity 94.9%; Pred. No. 3.5e-93;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IPKQYPIINFITAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 51  
DB 25 IPKQYPIINFITAGATVQSYNFIKAVRGLTGGADVRHDPVLPNRVGLPINQRFILV 84  
QY 52 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAPG 111  
DB 85 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAPG 144  
QY 112 GNYDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 171  
DB 145 GNYDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 204  
QY 172 FOYIEGEMTRIRYNRRS 189  
DB 205 FOYIEGEMTRIRYNRRS 222  
RESULT 12  
AAG78304  
ID AAG78304 standard; Protein; 565 AA.  
XX AAG78304;  
XX 27-NOV-2001 (first entry)  
XX Modified castor bean preproricin (SEQ ID 10).  
XX Castor bean plant; preproricin; ricin; A chain; B chain;  
KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
KW retroviral infection; anti-HIV; virucide; viral protease.  
XX Chimeric - Ricinus communis  
OS Chimeric - Human immunodeficiency virus type 2.  
XX Key Location/Qualifiers  
FH Peptide 1..24  
FT /label= Signal\_peptide  
FT Protein 25..565  
FT /label= Proridin  
FT /note= "Proridin consists of the ricin A chain, a linker  
FT peptide, and the ricin B chain. Proridin is  
FT proteolytically cleaved between the A chain and  
FT the linker to yield mature ricin"  
FT Protein 25..291  
FT /label= Ricin A chain  
FT /note= "N-glycosidase"  
FT Peptide 292..303  
FT /label= Linker\_peptide  
FT Cleavage-site 296..297

FT /label= HIV protease\_cleavage\_site  
 FT 304..565  
 FT /label= Ricin B chain  
 FT /note= "Galactose/N-acetylgalactosamine-binding lectin"  
 XX  
 XX WO200160393-A1.  
 XX  
 XX 23-AUG-2001.  
 XX  
 XX 15-FEB-2001; 2001WO-US05282.  
 XX  
 XX 16-FEB-2000; 2000US-0182759.  
 XX  
 XX (BECH-) BECHTEL BWXT IDAHO LLC.  
 XX  
 XX Keener WK, Ward TE;  
 XX  
 XX WPI: 2001-581908/65.  
 XX N-PSDB; AA164145.  
 XX  
 XX Novel composition comprising toxin e.g., ricin based antiviral compound  
 PT useful for treating viral infections such as human immunodeficiency  
 PT virus infection.  
 PT  
 PT  
 XX Example 1; Page 59-63; 66pp; English.  
 XX  
 XX The sequence relates to the amino acid sequence of a modified prepro-ricin  
 CC protein encoded by AA164145. The invention relates to a novel toxin  
 CC (e.g. ricin) based antiviral agent which is toxic to virus-infected  
 CC cells, but non-toxic to uninfected cells. The invention has anti-HIV and  
 CC virucide activities. The agent is able to enter all HIV susceptible  
 CC cells, and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell unless the cell is infected  
 CC with the HIV virus, where the viral protease activates it. Ricin's  
 CC mechanism of action is through inactivation of cellular ribosomes and  
 CC enhancement of binding of the antiviral agent to galactose residues on  
 CC cell surfaces, and its cellular internalisation. The invention is useful  
 CC for treating human immunodeficiency virus infection and other viral  
 CC infections, especially retroviral infections. The antiviral agent is  
 CC activated in viral particles or early-stage infected cells, killing the  
 CC cells upon infection and effectively preventing the integration of the  
 CC viral genome into the host genome thereby preventing the latency/rebound  
 CC problem.  
 XX Sequence 565 AA;  
 XX  
 XX Query Match 97.9%; Score 951; DB 22; Length 565;  
 XX Best Local Similarity 94.9%; Pred. No. 3.5e-93;  
 XX Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IPKQYPIINFITAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 DB 25 IPKQYPIINFITAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84  
 QY 52 ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 111  
 DB 85 ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 144  
 QY 112 GNYDRLEQLAGNLRENIELGNGPLEARISALYYSTGGTQLPTLARSFFIICIMISEAR 171  
 DB 145 GNYDRLEQLAGNLRENIELGNGPLEARISALYYSTGGTQLPTLARSFFIICIMISEAR 204  
 QY 172 FOYIEGEMTRIRNRRS 189  
 DB 205 FOYIEGEMTRIRNRRS 222  
 RESULT 13  
 AAP70326  
 ID AAP70326 standard; Protein; 576 AA.  
 XX  
 XX AAP70326;  
 XX

DT 25-MAR-2003 (updated)  
 DT 21-MAY-1991 (first entry)  
 XX  
 DE Sequence of Ricinus communis (castor bean) Ricin toxin  
 DE (RT or ricin) E precursor encoded by prt38.  
 XX  
 KW Lectin; toxin protein; cytotoxic; cytostatic; castor bean;  
 KW plant toxin.  
 XX  
 OS Ricinus communis.  
 XX  
 XX Location/Qualifiers  
 FH Key 1..35  
 FH Region /note= "leader"  
 FT 36..302  
 FT Region /note= "A-chain"  
 FT 315..576  
 FT Region /note= "B-chain"  
 FT  
 XX  
 PW EP237676-A.  
 XX  
 XX 23-SEP-1987.  
 PD  
 XX 13-NOV-1986; 86EP-0308877.  
 PF  
 XX 07-MAR-1986; 86US-0837583.  
 PR  
 XX (CETU ) CETUS CORP.  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Piatak M;  
 XX  
 XX WPI: 1987-265177/38.  
 DR N-PSDB; AA70526.  
 DR  
 XX New non-glycosylated ricin precursor and toxin etc. - are prepd.  
 PT by recombinant DNA procedures with specific isolation steps for  
 PT purer and soluble prods.  
 PT  
 XX Disclosure; Fig 14(1-2); 112pp; English.  
 XX  
 XX The full length sequences encoding ricin A (AA70520), ricin D  
 CC (AA70525) putative ricin E (AA70526) and RCA (AA70524) in precursor  
 CC form were obtained, using the messenger RNA to obtain a cDNA library, and  
 CC then probing the library to retrieve the desired cDNA inserts. The  
 CC library was probed using the 35-mer given in AA70514. Figure 4 (see  
 CC AA70520, AA70521, AA70522), shows the nucleotide sequences of three  
 CC plasmids containing cDNA inserts obtained by probing a cDNA library  
 CC for sequences encoding ricin B using the probe in AA70517. The cDNA  
 CC inserts can be placed into expression vectors. Site-directed  
 CC mutagenesis may be used to place an ATG start codon and a HindIII  
 CC site at the beginning of the mature protein. (see AA70518). The  
 CC coding sequences of the inserts can be ligated into expression  
 CC vectors containing the PhdA promoter-operator and leader sequence  
 CC (AA70523) and suitable retroregulators.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 XX Sequence 576 AA;  
 SQ  
 Query Match 97.9%; Score 951; DB 8; Length 576;  
 XX Best Local Similarity 94.9%; Pred. No. 3.6e-93;  
 XX Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 2 IPKQYPIINFITAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
 DB 36 IPKQYPIINFITAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 95  
 QY 52 ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 111  
 DB 96 ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 155  
 QY 112 GNYDRLEQLAGNLRENIELGNGPLEARISALYYSTGGTQLPTLARSFFIICIMISEAR 171



DB 156 GNYDRLEQLAGNLELNGPLEBAISALYYSTGGTQTLPTLARSFIICQMISEAAR 215

QY 172 FQYIEGEMRTRIRYNRRS 189

DB 216 FQYIEGEMRTRIRYNRRS 233

RESULT 14

AAW25787

ID AAW25787 standard; Protein; 576 AA.

XX AAW25787;

XX AAW25787;

DT 25-MAR-2003 (updated)

DT 27-MAR-1998 (first entry)

XX Castor bean ricin.

XX Ricin; cytotoxin; hybrid protein; cell delivery;

KW cell binding ligand; translocation domain; diphtheria toxin B';

KW interleukin-2; T-cell lymphoma; organ rejection; therapy.

XX Ricinus communis.

OS

XX Key Location/Qualifiers

FE Peptide 1..35

FT /label= Sig\_peptide

FT Protein 36..302

FT /label= A-domain

FT Peptide 303..314

FT /label= Linker

FT Domain 315..576

FT /label= B-domain

XX US5668255-A.

PN 16-SEP-1997.

PD

XX 04-AUG-1993; 93US-0102387.

PF

XX 27-JUN-1991; 91US-0722484.

PR 07-JUN-1984; 84US-0618199.

PR 25-APR-1985; 85US-0726808.

PR 07-JUN-1985; 85US-0742554.

PR 22-DEC-1989; 89US-0456095.

PR 14-JUN-1990; 90US-0538276.

PR 04-AUG-1993; 93US-0102387.

XX (SERA-) SERAGEN INC.

PA

XX Murphy JR;

PI

DR WPI; 1997-470103/43.

DR N-PSDB; AAT91638.

XX

XX New hybrid molecules for delivery of agents to cells - comprise a

PT binding domain of a cell binding ligand and a portion of a

PT translocation domain of a protein

XX

PS Example 4; Fig 11A-B; 30pp; English.

XX

CC DNA (see AAT91638) encoding the castorbean cytotoxin, ricin.

CC of the A-to-B linker peptide of ricin was used to construct a

CC ricin-diphtheria toxin B'-interleukin-2 gene that was expressed in

CC E. coli. The hybrid protein can be isolated and used to treat

CC conditions involving over-production of cells bearing IL2 receptors,

CC such as certain T-cell lymphomas and organ transplant rejection

CC crises. The hybrid inactivates ribosomes in cells bearing IL2

CC receptors, resulting in cessation of protein synthesis and death of

CC target cells. Claimed hybrid proteins comprise a translocation

CC domain and a cell binding domain from e.g. a hormone, growth factor

CC or polypeptide toxin. The hybrid molecules can be used for the

CC delivery of agents (e.g. therapeutic genes, toxins, detectable

CC labels) into cells. The use of a translocation mechanism ensures

CC that the hybrid will be effective in relatively low doses, since a

CC high proportion of the substance of interest will be taken into the

CC targeted cells. The hybrid molecules can be manufactured as a

CC single hybrid recombinant protein, permitting reproducibility,

CC consistency, and the precise control of composition.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 576 AA;

Query Match 97.9%; Score 951; DB 18; Length 576;

Best Local Similarity 94.9%; Pred No. 3.6e-93;

Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFETAGTAGVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51

DB 36 IFPKQYPIINFETAGTAGVQSYTNFIRAVRGRLTGDVDRHEIPVLPNRVGLPINQRFILV 95

QY 52 ELSNHAELSVTALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVQNRYYTFAFG 111

DB 96 ELSNHAELSVTALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHTLFTDVQNRYYTFAFG 155

QY 112 GNYDRLEQLAGNLELNGPLEBAISALYYSTGGTQTLPTLARSFIICQMISEAAR 171

DB 156 GNYDRLEQLAGNLELNGPLEBAISALYYSTGGTQTLPTLARSFIICQMISEAAR 215

QY 172 FQYIEGEMRTRIRYNRRS 189

DB 216 FQYIEGEMRTRIRYNRRS 233

RESULT 15

AAW55892

ID AAW55892 standard; Protein; 576 AA.

XX AAW55892;

AC

XX 15-FEB-2000 (first entry)

DT

XX Castor bean ricin toxin.

DE

XX Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria;

KW translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;

KW shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;

KW cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;

KW adipocyte; cancer; virus; infection; antibody.

XX Ricinus communis.

OS

XX US5965406-A.

PN

XX 12-OCT-1999.

PD

XX 07-JUN-1995; 95US-0488246.

PF

XX 04-AUG-1993; 93US-0102387.

PR 07-JUN-1984; 84US-0618199.

PR 27-JUN-1991; 91US-0722484.

PR 25-APR-1985; 85US-0726808.

PR 07-JUN-1985; 85US-0742554.

PR 22-DEC-1989; 89US-0456095.

PR 14-JUN-1990; 90US-0538276.

XX (SERA-) SERAGEN INC.

PA

XX Murphy JR;

PI

DR WPI; 1999-632431/54.

DR N-PSDB; AAZ30663.

XX Recombinant DNA molecule encoding a three part hybrid protein used in

PT the treatment of Aids and genetic deficiency diseases -

XX

PS Example 4; Fig 11; 31pp; English.

XX

XX The invention relates to a recombinant DNA molecule encoding a hybrid  
CC protein comprising three parts: (a) the first part comprises a portion  
CC of the binding domain of a cell-binding polypeptide ligand allowing the  
CC hybrid protein to bind to an animal cell; (b) the second part comprises  
CC a portion of a translocation domain of a naturally occurring protein  
CC selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera  
CC toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus  
CC toxin, which translocate the third part of the across the cytoplasmic  
CC membrane into the cytosol of the cell; and (c) the third part comprises  
CC a polypeptide entity to be introduced into the cell, which is non-native  
CC to the naturally occurring protein of (b). This sequence represents the  
CC Castor bean ricin toxin sequence for use in generating the hybrid of the  
CC invention. The hybrid molecule enables the direction of appropriate  
CC therapy to affected cells, allowing them to function properly and  
CC alleviate or cure the disease. The hybrid is especially used in treating  
CC genetic deficiency diseases, by delivering to affected cells an enzyme  
CC supplying the missing function, to supplementing cellular levels of a  
CC particular enzyme or a scarce precursor or cofactor, to directing toxins  
CC or other poisons to destroy particular cells (such as adipocytes, cancer  
CC cell, or virus infected-cells), to counteracting viral infections such as  
CC HIV, by introducing appropriate antibodies to viral proteins. It is also  
CC involved in the process of getting non-therapeutic substances such as  
CC detectable labels into cells.

XX

SQ Sequence 576 AA;

Query Match		97.9%;	Score 951;	DB 20;	Length 576;
Best Local Similarity		94.9%;	Pred. No. 3.6e-93;		
Matches 188;		Conservative 0;	Mismatches 0;	Indels 10;	Gaps 1;
QY	2	IFPKQYPIINF	TAGATVQSYTNFIRAVRGLT-----	VLPNRVGLP	INQRFILV 51
Db	36	IFPKQYPIINF	TAGATVQSYTNFIRAVRGLT	TGADVREH	IPVLPNRVGLP
QY	52	ELSNHAELSV	TALDVNTNAYVVGVRAGNSGAYFFHPDNOQEDAEAI	THLFTD	VQNRYYTFAFG 111
Db	96	ELSNHAELSV	TALDVNTNAYVVGVRAGNSGAYFFHPDNOQEDAEAI	THLFTD	VQNRYYTFAFG 155
QY	112	GNVDRLEQL	AGNLRNIELNGNPLEEALISALYYSTGGTQLP	TLARSF	IIQIMISEAAR 171
Db	156	GNVDRLEQL	AGNLRNIELNGNPLEEALISALYYSTGGTQLP	TLARSF	IIQIMISEAAR 215
QY	172	FOYIEGEM	TRIRYNRRS 189		
Db	216	FOYIEGEM	TRIRYNRRS 233		

Search completed: February 10, 2004, 16:22:27

Job time : 33.2239 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 10.9281 Seconds  
(without alignments)  
731.761 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MIFPKQPIINFTAGATVQ.....ARFQVIEGEMTRIRYNRRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/PCTUS-COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	956	98.5	268	2	US-08-356-786-8
2	956	98.5	534	2	US-08-356-786-10
3	951	97.9	267	1	US-07-901-707-1
4	951	97.9	267	1	US-07-988-430-1
5	951	97.9	267	1	US-08-425-336-1
6	951	97.9	267	1	US-08-488-1138-1
7	951	97.9	267	2	US-08-477-484B-1
8	951	97.9	267	2	US-08-646-360-1
9	951	97.9	267	3	US-08-839-765-1
10	951	97.9	267	3	US-09-136-389-1
11	951	97.9	267	4	US-09-610-838-1
12	951	97.9	267	5	PCT-US92-09487-1
13	951	97.9	290	1	US-08-378-761A-27
14	951	97.9	290	1	US-08-485-286-27
15	951	97.9	290	6	5248606-4
16	941	96.9	267	1	US-08-218-303-16
17	941	96.9	267	2	US-08-338-793D-61
18	941	96.9	267	4	US-09-538-873-1
19	861.5	88.7	540	1	US-08-378-761A-77
20	861.5	88.7	540	1	US-08-485-286-77
21	336	34.6	247	1	US-08-488-113B-6
22	336	34.6	247	1	US-08-477-484B-6
23	336	34.6	247	2	US-08-646-360-6
24	336	34.6	247	3	US-08-839-765-6
25	336	34.6	247	3	US-09-136-389-6
26	336	34.6	247	4	US-09-610-838-6
27	336	34.6	267	1	US-08-378-761A-74

28	336	34.6	267	1	US-08-485-286-74
29	336	34.6	289	1	US-07-923-692C-4
30	336	34.6	289	1	US-08-184-237-4
31	336	34.6	289	2	US-08-482-920-4
32	336	34.6	289	3	US-08-484-341-4
33	336	34.6	289	3	US-08-483-502-4
34	336	34.6	289	4	US-09-726-651A-4
35	320.5	33.0	282	1	US-08-324-301-15
36	310.5	32.0	255	1	US-07-901-707-6
37	310.5	32.0	255	1	US-07-988-430-6
38	310.5	32.0	255	1	US-08-425-336-6
39	310.5	32.0	255	5	PCT-US92-09487-6
40	305	31.4	248	3	US-08-902-486-7
41	305	31.4	290	1	US-08-245-754A-2
42	305	31.4	290	2	US-08-597-731-2
43	305	31.4	496	3	US-08-902-486-15
44	303.5	31.3	250	1	US-08-378-761A-71
45	303.5	31.3	250	1	US-08-485-286-71

ALIGNMENTS

RESULT 1  
US-08-356-786-8  
; Sequence 8, Application US/08356786  
; Patent No. 5877305  
; GENERAL INFORMATION:  
; APPLICANT: Huston, James S.  
; APPLICANT: Oppermann, Hermann  
; APPLICANT: Houston, L. L.  
; APPLICANT: Ring, David B.  
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
; TITLE OF INVENTION: Marker  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,786  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/831,967  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 268 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-356-786-8

Query Match 98.5%; Score 956; DB 2; Length 268;  
Best Local Similarity 95.0%; Pred. NO. 7.1e-103;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50  
Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFIL 60  
QY 51 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTFAG 110  
Db 61 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTFAG 120  
QY 111 GGNDRLEQLAGNLRNENIELGNGLPLEEALSALYYSTGCTQLPTLARSFIICQMISEAA 170  
Db 121 GGNDRLEQLAGNLRNENIELGNGLPLEEALSALYYSTGCTQLPTLARSFIICQMISEAA 180  
QY 171 RFQYIEGEMTRIRYNRS 189  
Db 181 RFQYIEGEMTRIRYNRS 199

## RESULT 2

US-08-356-786-10  
; Sequence 10, Application US/08356786  
; Patent No. 5877305  
; GENERAL INFORMATION:  
; APPLICANT: Huston, James S.  
; APPLICANT: Oppermann, Hermann  
; APPLICANT: Houston, L. L.  
; APPLICANT: Ring, David B.  
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
; MARKER  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,786  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/831,967  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 534 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-356-786-10

Query Match 98.5%; Score 956; DB 2; Length 534;  
Best Local Similarity 95.0%; Pred. No. 2e-102;  
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50  
Db 3 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFIL 62  
QY 51 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTFAG 110

Db 63 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTFAG 122  
QY 111 GGNDRLEQLAGNLRNENIELGNGLPLEEALSALYYSTGCTQLPTLARSFIICQMISEAA 170  
Db 123 GGNDRLEQLAGNLRNENIELGNGLPLEEALSALYYSTGCTQLPTLARSFIICQMISEAA 182  
QY 171 RFQYIEGEMTRIRYNRS 189  
Db 183 RFQYIEGEMTRIRYNRS 201

## RESULT 3

US-07-901-707-1  
; Sequence 1, Application US/07901707  
; Patent No. 5376546  
; GENERAL INFORMATION:  
; APPLICANT: Bernhard, Susan L.  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Steve F.  
; APPLICANT: Lane, Julie A.  
; TITLE OF INVENTION: Materials Comprising and Methods of  
; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins  
; NUMBER OF SEQUENCES: 57  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell,  
; STREET: Two First National Plaza, 20 South Clark  
; STREET: Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/901,707  
; FILING DATE: 19920619  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5376546and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27129/30910  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-5750  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-901-707-1

Query Match 97.9%; Score 951; DB 1; Length 267;  
Best Local Similarity 94.9%; Pred. No. 2.7e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFIL 60  
QY 52 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTFAG 111  
Db 61 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTFAG 120  
QY 112 GGNDRLEQLAGNLRNENIELGNGLPLEEALSALYYSTGCTQLPTLARSFIICQMISEAA 171

Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRS 189  
Db 181 FOYIEGEMTRIRYNRS 198

RESULT 4  
US-07-988-430-1  
; Sequence 1, Application US/07988430  
; Patent No. 5416202  
; GENERAL INFORMATION:  
; APPLICANT: Bernhard, Susan L.  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Lane, Julie A.  
; APPLICANT: Lei, Shau-Ping  
; TITLE OF INVENTION: Materials Comprising and Methods of  
; TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins  
; NUMBER OF SEQUENCES: 101  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; STREET: Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/988,430  
; FILING DATE: 19921209  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5416202and, Greta E.  
; REGISTRATION NUMBER: 35302  
; REFERENCE/DOCKET NUMBER: 31133  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-988-430-1

Query Match 97.9%; Score 951; DB 1; Length 267;  
Best Local Similarity 94.9%; Pred. No. 2.7e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLIV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINORFLIV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120

QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFIIQIMISEAAR 171  
Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRS 189  
Db 181 FOYIEGEMTRIRYNRS 198

RESULT 5  
US-08-425-336-1  
; Sequence 1, Application US/08425336  
; Patent No. 5621083  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; ADDRESSEE: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/425,336  
; FILING DATE: 18-APR-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/064,691  
; FILING DATE: 12-MAY-1993  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Thomas C.  
; REGISTRATION NUMBER: P-36,989  
; REFERENCE/DOCKET NUMBER: 31394  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-425-336-1

Query Match 97.9%; Score 951; DB 1; Length 267;  
Best Local Similarity 94.9%; Pred. No. 2.7e-102;  
Matches 186; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLIV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINORFLIV 60  
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQLPTLARSFIIQIMISEAAR 171

Db 121 GNYDRLEQAGNLRNIEIENGPLEEASALVYSTGGTQPLTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRS 189  
Db 181 FOYIEGEMTRIRYNRS 198  
RESULT 6  
US-08-488-113B-1  
; Sequence 1, Application US/08488113B  
; Patent No. 5744580  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION NUMBER: US/08/488,113B  
; FILING DATE: 18-APR-1995  
; APPLICATION NUMBER: US 08/425,336  
; PRIOR APPLICATION DATA:  
; FILING DATE: 12-MAY-1993  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; FILING DATE: 09-DEC-1992  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-488-113B-1  
Query Match 97.9%; Score 951; DB 1; Length 267;  
Best Local Similarity 94.9%; Pred. No. 2.7e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFATTAGATVQSYTNFIRAVRGLTTCGADVRHEIPVLPNRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLETDYQNRVTFAPG 111  
Db 61 ELSNHAELSVTLALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLETDYQNRVTFAPG 120  
QY 112 GNYDRLEQAGNLRNIEIENGPLEEASALVYSTGGTQPLTLARSFIIQIMISEAAR 171  
Db 121 GNYDRLEQAGNLRNIEIENGPLEEASALVYSTGGTQPLTLARSFIIQIMISEAAR 180  
QY 172 FOYIEGEMTRIRYNRS 189  
Db 181 FOYIEGEMTRIRYNRS 198  
RESULT 7  
US-08-477-484B-1  
; Sequence 1, Application US/08477484B  
; Patent No. 5756699  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION NUMBER: US/08/477,484B  
; FILING DATE: 18-APR-1995  
; APPLICATION NUMBER: US 08/425,336  
; PRIOR APPLICATION DATA:  
; FILING DATE: 12-MAY-1993  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; FILING DATE: 09-DEC-1992  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; TELEX: 650 388-1248  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-477-484B-1  
Query Match 97.9%; Score 951; DB 1; Length 267;

Best Local Similarity 94.9%; Pred. No. 2.7e-102; Indels 10; Gaps 1;  
Matches 188; Conservative 0; Mismatches 0;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVDRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELS VTLALDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELS VTLALDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEALISALYYVSTGTQTLPTLARSFIICIMISEAAR 171  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEALISALYYVSTGTQTLPTLARSFIICIMISEAAR 180  
QY 172 FOYIEGEMRTRIRYNNRS 189  
Db 181 FOYIEGEMRTRIRYNNRS 198

RESULT 8  
US-08-646-360-1  
; Sequence 1, Application US/08646360  
; Patent No. 5837491  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/646,360  
; FILING DATE: 13-MAY-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05348  
; FILING DATE: 12-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 200-70.P4  
; TELEPHONE: 312/707-8889  
; TELEFAX: 312/707-9155  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-646-360-1

Query Match 97.9%; Score 951; DB 2; Length 267;  
Best Local Similarity 94.9%; Pred. No. 2.7e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51  
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVDRHEIPVLPNRVGLPINQRFILV 60  
QY 52 ELSNHAELS VTLALDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111  
Db 61 ELSNHAELS VTLALDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120  
QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEALISALYYVSTGTQTLPTLARSFIICIMISEAAR 171  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEALISALYYVSTGTQTLPTLARSFIICIMISEAAR 180  
QY 172 FOYIEGEMRTRIRYNNRS 189  
Db 181 FOYIEGEMRTRIRYNNRS 198

RESULT 9  
US-08-839-765-1  
; Sequence 1, Application US/08839765  
; Patent No. 6146631  
; GENERAL INFORMATION:  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Stephen F.  
; APPLICANT: Studnika, Gary M.  
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 169  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
; STREET: 500 West Madison Street, 34th floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60661  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/839,765  
; FILING DATE: 15-APR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/425,336  
; FILING DATE: 18-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/064,691  
; FILING DATE: 12-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/988,430  
; FILING DATE: 09-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,707  
; FILING DATE: 19-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/787,567  
; FILING DATE: 04-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McNicholas, Janet M.  
; REGISTRATION NUMBER: 32,918  
; REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-765-1

Query Match 97.9%; Score 951; DB 3; Length 267;  
Best Local Similarity 94.9%; Pred. No. 2.7e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLT-----VLPNRVGLPINQRFILV 51  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTGADVRHEIPVLPNRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 120

QY 112 GNYDRLEQLAGNLENLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 171  
DB 121 GNYDRLEQLAGNLENLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 180

QY 172 FOYIEGEMTRIRYNRS 189  
DB 181 FOYIEGEMTRIRYNRS 198

RESULT 10  
US-09-136-389-1  
Sequence 1, Application US/09136389  
Patent No. 6146850  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/09/136,389  
FILING DATE: 06-JUL-2000  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-136-389-1

Query Match 97.9%; Score 951; DB 3; Length 267;  
Best Local Similarity 94.9%; Pred. No. 2.7e-102;  
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLT-----VLPNRVGLPINQRFILV 51  
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVGRGLTGADVRHEIPVLPNRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111  
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 120

QY 112 GNYDRLEQLAGNLENLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 171  
DB 121 GNYDRLEQLAGNLENLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 180

QY 172 FOYIEGEMTRIRYNRS 189  
DB 181 FOYIEGEMTRIRYNRS 198

RESULT 11  
US-09-610-838-1  
Sequence 1, Application US/09610838  
Patent No. 6376217  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
TITLE OF INVENTION: Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/09/610,838  
FILING DATE: 06-JUL-2000  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,389  
FILING DATE: 18-AUG-1998  
APPLICATION NUMBER: 08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348



```

; FILING DATE: 12-MAY-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-610-838-1

```

```

Query Match 97.9%; Score 951; DB 4; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IPFKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 1 IPFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVHRHIEPVLNVRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEALHTLFTDVQNYTFAG 111
DB 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEALHTLFTDVQNYTFAG 120

QY 112 GNYDRLEQLAGNLENIELGNGPLEEALISALYYSTGTQTLPTLARSFIIQIMISEAAR 171
DB 121 GNYDRLEQLAGNLENIELGNGPLEEALISALYYSTGTQTLPTLARSFIIQIMISEAAR 180

QY 172 FQYIEGEMTRIRYNRRS 189
DB 181 FQYIEGEMTRIRYNRRS 198

```

```

RESULT 12
PCT-US92-09487-1
; Sequence 1, Application PC/TUS9209487
; GENERAL INFORMATION:
; APPLICANT: Bernhardt, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Lane, Julie A.
; APPLICANT: Lei, Shau-Ping
; TITLE OF INVENTION: Materials Comprising and Methods of
; PREPARATION AND USE FOR RIBOSOME-INACTIVATING PROTEINS
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Hicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09487
; FILING DATE: 19921104
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35302
; REFERENCE/DOCKET NUMBER: 31133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-09487-1

```

```

Query Match 97.9%; Score 951; DB 5; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IPFKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 1 IPFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVHRHIEPVLNVRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEALHTLFTDVQNYTFAG 111
DB 61 ELSNHAELSVTLADVTNAYVVGVRAGNSAYFFHPDNOEDAEALHTLFTDVQNYTFAG 120

QY 112 GNYDRLEQLAGNLENIELGNGPLEEALISALYYSTGTQTLPTLARSFIIQIMISEAAR 171
DB 121 GNYDRLEQLAGNLENIELGNGPLEEALISALYYSTGTQTLPTLARSFIIQIMISEAAR 180

QY 172 FQYIEGEMTRIRYNRRS 189
DB 181 FQYIEGEMTRIRYNRRS 198

```

```

RESULT 13
US-08-378-761A-27
; Sequence 27, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; PREPARATION AND USE FOR RIBOSOME-INACTIVATING PROTEINS
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/378,761A

FILING DATE: 26-JAN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 290 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-378-761A-27

Query Match 97.9%; Score 951; DB 1; Length 290;

Best Local Similarity 94.9%; Pred. No. 3e-102;

Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```
QY      2  IFPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51
Db      25  IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84
QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111
Db      85  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 144
QY     112  GNYDRLEQLAGNLRNIELNGPLLEAEISALYYSTGGTQPLTLARSFIIICQMISEAAR 171
Db     145  GNYDRLEQLAGNLRNIELNGPLLEAEISALYYSTGGTQPLTLARSFIIICQMISEAAR 204
QY     172  FOYIEGEMRTRIRYNRRS 189
Db     205  FOYIEGEMRTRIRYNRRS 222
```

RESULT 14

US-08-485-286-27

Sequence 27, Application US/08485286

Patent No. 5646026

Patent No. 5646026 5646119

GENERAL INFORMATION:

APPLICANT: WALSH, TERENCE A

APPLICANT: HEY, TIMOTHY D

APPLICANT: MORGAN, ALICE ER

TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE

TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF

TITLE OF INVENTION: USING

NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: ANDREA T. BORUCKI

STREET: 9330 ZIONSVILLE ROAD

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: US

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,286

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/378761

FILING DATE: 26-JAN-1995

Search completed: February 10, 2004, 16:29:32

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 290 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-485-286-27

Query Match 97.9%; Score 951; DB 1; Length 290;

Best Local Similarity 94.9%; Pred. No. 3e-102;

Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```
QY      2  IFPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51
Db      25  IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84
QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111
Db      85  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 144
QY     112  GNYDRLEQLAGNLRNIELNGPLLEAEISALYYSTGGTQPLTLARSFIIICQMISEAAR 171
Db     145  GNYDRLEQLAGNLRNIELNGPLLEAEISALYYSTGGTQPLTLARSFIIICQMISEAAR 204
QY     172  FOYIEGEMRTRIRYNRRS 189
Db     205  FOYIEGEMRTRIRYNRRS 222
```

RESULT 15

5248606-4

Patent No. 5248606

APPLICANT: WALSH, TERENCE A; HEY, TIMOTHY D; MORGAN,

ALICE E.R.

TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND

ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATION

NUMBER OF SEQUENCES: 49

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/535,636

SEQ ID NO: 4

FILING DATE: 11-JUN-1990

LENGTH: 290

5248606-4

Query Match

Best Local Similarity 97.9%; Score 951; DB 6; Length 290;

Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```
QY      2  IFPKQYPIINFTTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51
Db      25  IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84
QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111
Db      85  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 144
QY     112  GNYDRLEQLAGNLRNIELNGPLLEAEISALYYSTGGTQPLTLARSFIIICQMISEAAR 171
Db     145  GNYDRLEQLAGNLRNIELNGPLLEAEISALYYSTGGTQPLTLARSFIIICQMISEAAR 204
QY     172  FOYIEGEMRTRIRYNRRS 189
Db     205  FOYIEGEMRTRIRYNRRS 222
```

Sun Feb 15 07:30:00 2004

us-10-083-336a-6.rai

Page 9

Job time : 10.9281 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:26:46 ; Search time 24.3781 Seconds  
(without alignment)  
1623.314 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971  
Sequence: 1 MIFPKQYPIINFTAGATVQ.....ARFOYIEGEMTRIVNRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	971	100.0	189	12	US-10-083-336A-6
2	971	100.0	190	12	US-10-083-336A-11
3	966	99.5	188	12	US-10-083-336A-4
4	956	98.5	188	12	US-10-083-336A-8
5	956	98.5	199	12	US-10-083-336A-5
6	956	98.5	200	12	US-10-083-336A-10
7	951	97.9	198	12	US-10-083-336A-3
8	951	97.9	267	12	US-10-127-890-1
9	951	97.9	576	12	US-10-083-336A-1
10	941	96.9	198	12	US-10-083-336A-7
11	941	96.9	267	12	US-10-282-935-1
12	941	96.9	267	12	US-10-440-796-1
13	939.5	96.8	185	12	US-10-083-336A-9
14	679	69.9	179	12	US-10-083-336A-2
15	336	34.6	247	10	US-09-792-793A-39

16	336	34.6	247	12	US-10-127-890-6
17	336	34.6	247	12	US-10-375-209A-39
18	336	34.6	289	12	US-10-280-679B-4
19	305	31.4	247	10	US-09-792-793A-34
20	305	31.4	247	12	US-10-375-209A-34
21	297.5	30.6	251	12	US-10-282-935-3
22	297.5	30.6	251	12	US-10-440-796-3
23	273	28.1	263	12	US-10-127-890-7
24	272	28.0	263	12	US-10-127-890-4
25	257.5	26.5	248	12	US-10-127-890-5
26	254	26.2	252	9	US-09-347-064-2
27	254	26.2	252	9	US-09-347-064-8
28	245.5	25.3	251	12	US-10-127-890-107
29	244.5	25.2	251	12	US-10-127-890-106
30	244.5	25.2	251	12	US-10-127-890-110
31	244.5	25.2	251	12	US-10-127-890-111
32	243.5	25.1	251	9	US-09-765-527-247
33	243.5	25.1	251	12	US-10-127-890-2
34	243.5	25.1	251	12	US-10-127-890-99
35	243.5	25.1	251	12	US-10-127-890-100
36	243.5	25.1	251	12	US-10-127-890-101
37	243.5	25.1	251	12	US-10-127-890-102
38	243.5	25.1	251	12	US-10-127-890-103
39	243.5	25.1	251	12	US-10-127-890-104
40	243.5	25.1	251	12	US-10-127-890-105
41	243.5	25.1	316	12	US-10-074-596-1
42	243.5	25.1	507	12	US-10-074-596-11
43	242.5	25.0	251	12	US-10-127-890-109
44	242.5	25.0	293	9	US-09-765-527-259
45	242.5	25.0	309	9	US-09-765-527-253

ALIGNMENTS

RESULT 1  
US-10-083-336A-6  
; Sequence 6, Application US/10083336A  
; Publication No. US20030181665A1  
; GENERAL INFORMATION:  
; APPLICANT: Olson, Mark A  
; APPLICANT: Millard, Charles B  
; APPLICANT: Byrne, Michael P  
; APPLICANT: Wannenmacher, Robert W  
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof  
; FILE REFERENCE: P67452USO (RIID 01.58)  
; CURRENT APPLICATION NUMBER: US/10/083,336A  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 189  
; TYPE: PRT  
; ORGANISM: Ricinus communis  
US-10-083-336A-6

Query Match	100.0%	Score 971;	DB 12;	Length 189;
Best Local Similarity	100.0%	Pred. No. 2.7e-103;		
Matches 189;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MIFPKQYPIINFTAGATVQSYTNFIRAVRGRLTVLPNRVGLPINQRFILVELSNHAELS	60	
Db	1	MIFPKQYPIINFTAGATVQSYTNFIRAVRGRLTVLPNRVGLPINQRFILVELSNHAELS	60	
QY	61	VTALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNYTFAFGNYDLEQL	120	
Db	61	VTALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNYTFAFGNYDLEQL	120	
QY	121	AGNLRENIELGNPLEEASALYYSTGTQTPTLARSFIIICIMTSEARFOYIEGEMR	180	
Db	121	AGNLRENIELGNPLEEASALYYSTGTQTPTLARSFIIICIMTSEARFOYIEGEMR	180	
QY	181	TRIVNRS	189	

```
Db 181 TRIRYRNS 189
|||||
1 IPKQYPIINFTAGATVQSYTNFIRAVRGRLTLPNVRGLPINQRFILVELSNHAELSV 60

RESULT 2
US-10-083-336A-11
; Sequence 11, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnes, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-11

Query Match 100.0%; Score 971; DB 12; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.7e-103;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIPKQYPIINFTAGATVQSYTNFIRAVRGRLTLPNVRGLPINQRFILVELSNHAELS 60
Db 1 MIPKQYPIINFTAGATVQSYTNFIRAVRGRLTLPNVRGLPINQRFILVELSNHAELS 60
QY 61 VTALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFGGYDRLEQL 120
Db 61 VTALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFGGYDRLEQL 120
QY 121 AGNREINELGNGLPBEAISALYYSTGGTQLPPLARSFIIICIMISEAARFOYIEGEMR 180
Db 121 AGNREINELGNGLPBEAISALYYSTGGTQLPPLARSFIIICIMISEAARFOYIEGEMR 180
QY 181 TRIRYRNS 189
Db 181 TRIRYRNS 189

RESULT 3
US-10-083-336A-4
; Sequence 4, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnes, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-4

Query Match 99.5%; Score 966; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 9.9e-103;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IPKQYPIINFTAGATVQSYTNFIRAVRGRLTLPNVRGLPINQRFILVELSNHAELSV 61
|||||

Db 181 TRIRYRNS 189
|||||
1 IPKQYPIINFTAGATVQSYTNFIRAVRGRLTLPNVRGLPINQRFILVELSNHAELSV 60

62 TLALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFGGYDRLEQLA 121
61 TLALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFGGYDRLEQLA 120
122 GNREINELGNGLPBEAISALYYSTGGTQLPPLARSFIIICIMISEAARFOYIEGEMRT 181
121 GNREINELGNGLPBEAISALYYSTGGTQLPPLARSFIIICIMISEAARFOYIEGEMRT 180
182 RIRYRNS 189
181 RIRYRNS 188

RESULT 4
US-10-083-336A-8
; Sequence 8, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnes, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-8

Query Match 98.5%; Score 956; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.4e-101;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PKQYPIINFTAGATVQSYTNFIRAVRGRLTLPNVRGLPINQRFILVELSNHAELSVTL 63
Db 3 PKQYPIINFTAGATVQSYTNFIRAVRGRLTLPNVRGLPINQRFILVELSNHAELSVTL 62
QY 64 ALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFGGYDRLEQLAGN 123
Db 63 ALDVTNAYVGYRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFGGYDRLEQLAGN 122
QY 124 LRENIEGNGLPBEAISALYYSTGGTQLPPLARSFIIICIMISEAARFOYIEGEMTRI 183
Db 123 LRENIEGNGLPBEAISALYYSTGGTQLPPLARSFIIICIMISEAARFOYIEGEMTRI 182
QY 184 RYRNS 189
Db 183 RYRNS 188

RESULT 5
US-10-083-336A-5
; Sequence 5, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnes, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
```

```

; LENGTH: 199
; TYPE: PRT
; ORGANISM: Ricinus communis
; US-10-083-336A-5

Query Match      98.5%; Score 956; DB 12; Length 199;
Best Local Similarity 95.0%; Pred. No. 1.5e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50
DB 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFIL 60
QY 51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNYTFAP 110
DB 61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNYTFAP 120
QY 111 GGNVDRLEQLAGNLRNENIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAA 170
DB 121 GGNVDRLEQLAGNLRNENIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAA 180
QY 171 RFQYIEGEMTRIRYNRRS 189
DB 181 RFQYIEGEMTRIRYNRRS 199

RESULT 6
US-10-083-336A-10
; Sequence 10, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnes, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 10
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Ricinus communis
; US-10-083-336A-10

Query Match      98.5%; Score 956; DB 12; Length 200;
Best Local Similarity 95.0%; Pred. No. 1.5e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50
DB 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFIL 60
QY 51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNYTFAP 110
DB 61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNYTFAP 120
QY 111 GGNVDRLEQLAGNLRNENIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAA 170
DB 121 GGNVDRLEQLAGNLRNENIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAA 180
QY 171 RFQYIEGEMTRIRYNRRS 189
DB 181 RFQYIEGEMTRIRYNRRS 199

RESULT 7
US-10-083-336A-3
; Sequence 3, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnes, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Ricinus communis
; US-10-083-336A-3

Query Match      97.9%; Score 951; DB 12; Length 198;
Best Local Similarity 94.9%; Pred. No. 5.7e-101;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGADVRHEIPVLPNRVGLPINQRFILV 60
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNYTFAPG 111
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNYTFAPG 120
QY 112 GNVDRLEQLAGNLRNENIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAA 171
DB 121 GNVDRLEQLAGNLRNENIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAA 180
QY 172 FOYIEGEMTRIRYNRRS 189
DB 181 FOYIEGEMTRIRYNRRS 198

RESULT 8
US-10-127-890-1
; Sequence 1, Application US/10127890
; Publication No. US20030166196A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992

```

```

96  ELSNHAEISVTLALDVTNAYVGVYRAGNSAYFFHPDNOQEDAEAI THLFTDVQNRYYTFAFG 155
112 GNYDRLEQIAGNLRNIELGNGLPLEEASALYYSTGGTQLP T LARSFFI C IOMISEAAR 171
156 GNYDRLEQIAGNLRNIELGNGLPLEEASALYYSTGGTQLP T LARSFFI C IOMISEAAR 215
172 FOYIEGEMRTRIRYNRRS 189
216 FOYIEGEMRTRIRYNRRS 233

RESULT 10
US-10-083-336A-7
; Sequence 7, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-7

Query Match 96.9%; Score 941; DB 12; Length 198;
Best Local Similarity 94.9%; Pred. No. 8e-100;
Matches 186; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 4 PKQYPIINFATTAGATVQSYTNFTRAVRGRLT-----VLPNRVGLPINQRFILVEL 53
Db 3 PKQYPIINFATTAGATVQSYTNFTRAVRGRLT-----VLPNRVGLPINQRFILVEL 62
QY 54 SNHAEISVTLALDVTNAYVGVYRAGNSAYFFHPDNOQEDAEAI THLFTDVQNRYYTFAFGN 113
Db 63 SNHAEISVTLALDVTNAYVGVYRAGNSAYFFHPDNOQEDAEAI THLFTDVQNRYYTFAFGN 122
QY 114 YDRLEQIAGNLRNIELGNGLPLEEASALYYSTGGTQLP T LARSFFI C IOMISEAARFQ 173
Db 123 YDRLEQIAGNLRNIELGNGLPLEEASALYYSTGGTQLP T LARSFFI C IOMISEAARFQ 182
QY 174 YIEGEMRTRIRYNRRS 189
Db 183 YIEGEMRTRIRYNRRS 198

RESULT 11
US-10-282-935-1
; Sequence 1, Application US/10282935
; Publication No. US20030143193A1
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETTIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF
; FILE REFERENCE: UTSD:8840S
; CURRENT APPLICATION NUMBER: US/10/282,935
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 09/538,873
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/126,826
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1

```





Search completed: February 10, 2004, 16:53:53  
Job time : 25.3781 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:35 ; Search time 10.2276 Seconds  
(without alignments)  
1777.145 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MIFPKQYPIINFITAGATVQ.....ARFOYIEGEMTRIRYNRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: Pirl1.\*

2: Pirl2.\*

3: Pirl3.\*

4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	951	97.9	576	1	RLCSD
2	861.5	88.7	564	1	ricin D precursor
3	336	34.6	289	1	agglutinin precurs
4	327.5	33.7	528	2	rRNA N-glycosidase
5	327.5	33.7	562	2	abrin-d precursor
6	326	33.6	247	2	abrin-c precursor
7	326	33.6	247	2	karasurin-B - Mongol
8	326	33.6	247	2	karasurin C - Tric
9	323	33.3	527	2	abrin-b precursor
10	310.5	32.0	251	2	abrin (clone 7.2)
11	303.5	31.3	528	1	abrin-a precursor
12	300.5	30.9	278	2	beta-luffin - smoo
13	293.5	30.2	250	2	luffin-b - smooth
14	274.5	28.3	277	2	rRNA N-glycosidase
15	273	28.1	286	1	rRNA N-glycosidase
16	273	28.1	570	2	agglutinin I precu
17	272	28.0	286	2	rRNA N-glycosidase
18	270	27.8	245	2	rRNA N-glycosidase
19	265	27.3	286	2	rRNA N-glycosidase
20	264	27.2	254	2	mistletoe lectin I
21	243.5	25.1	316	2	rRNA N-glycosidase
22	186.5	19.2	294	2	rRNA N-glycosidase
23	174	17.9	278	2	rRNA N-glycosidase
24	170	17.5	313	2	rRNA N-glycosidase
25	168.5	17.4	261	2	antiviral protein
26	148.5	15.3	289	2	rRNA N-glycosidase
27	138	14.2	272	2	betavulgin - beet
28	131	13.5	253	2	rRNA N-glycosidase
29	127	13.1	253	2	rRNA N-glycosidase

#### ALIGNMENTS

##### RESULT 1

##### RLCSD

ricin D precursor - castor bean

N:Contains: rRNA N-glycosidase (EC 3.2.2.22)

C:Species: Ricinus communis (castor bean)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999

C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.P.

Nucleic Acids Res. 13, 8019-8033, 1985

A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A:Reference number: A24041; MUID:86067214; PMID:2999712

A:Accession: A24041

A:Molecule type: DNA

A:Residues: 1-576 <HAL>

A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R:Tregear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin gene

A:Reference number: S20513; MUID:92163016; PMID:1371405

A:Accession: S20513

A:Molecule type: DNA

A:Residues: 1-576 <TRE>

A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R:Lamb, F.I.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A:Title: Nucleotide sequence of cloned cDNA coding for preproricin.

A:Reference number: A24614; MUID:85179479; PMID:3838723

A:Accession: A24614

A:Molecule type: mRNA

A:Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>

A:Cross-references: GB:X02388; NID:g21077; PIDN:CAA26230.1; PID:g21078

R:Yoshitake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A:Title: Isolation and sequences of peptic peptides, and the complete sequence of Ile cha

A:Reference number: A03372

A:Accession: A03372

A:Molecule type: protein

A:Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 272-283, 'L', 285-288, 290-302 <YOS>

A:Note: this paper cites the others in the series providing experimental details for the

R:Araki, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1985

A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan

A:Reference number: A24010

A:Accession: A24010

A:Molecule type: protein

A:Residues: 315-383, 'PS', 386-576 <ARA>

R:Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A:Title: Primary structure of Ala chain of ricin D.

A:Reference number: A03374

A:Accession: A03374

A:Molecule type: protein

A;Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405-551, 'V', 553-572, 'E', 529-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F' <FUN>  
 A;Note: This paper, one of a series, summarizes the experimental details for the determination of the primary sequence of Ricinus communis agglutinin.  
 R;Ready, M.P.; Kim, Y.; Robertus, J.D.  
 Proteins 10, 270-278, 1991  
 A;Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism of ricin toxicity.  
 A;Reference number: A48237; PMID:91352006; PMID:1881883  
 A;Contents: annotation; active site  
 R;Robertus, J.D.  
 Proteins 10, 260-269, 1991  
 A;Title: Structure of ricin B-chain at 2.5 angstrom resolution.  
 A;Reference number: A48238; PMID:91352005; PMID:1881882  
 A;Contents: annotation; X-ray crystallography, 2.5 angstroms  
 R;Katzin, B.J.; Collins, E.J.; Robertus, J.D.  
 Proteins 10, 251-259, 1991  
 A;Title: Structure of ricin A-chain at 2.5 angstroms.  
 A;Reference number: A48239; PMID:91352004; PMID:1881881  
 A;Contents: annotation; X-ray crystallography, 2.5 angstroms  
 C;Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which is inactive. The A chain inhibits protein synthesis; it inactivates the 60S ribosomal subunit of the cell of the A chain; B chains are also responsible for cell agglutination (lectin).  
 C;Comment: This protein is cytotoxic and very poisonous to animals.  
 C;Superfamily: ricin; rRNA N-glycosidase homology  
 C;Keywords: duplication; glycoprotein; hydrolase; lectin; RNA binding; seed  
 F;1-35/Domain: signal sequence #status predicted <SIG>  
 F;36-302/Product: ricin D chain A #status experimental <ACH>  
 F;46-293/Domain: rRNA N-glycosidase homology <RNG>  
 F;315-576/Product: ricin D chain B #status experimental <BCH>  
 F;331-373, 374-414, 417-455, 462-497, 501-540, 543-576/Region: 40-residue repeats  
 F;45, 409, 449/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;115, 158, 243, 244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F;212/Active site: Glu #status experimental  
 F;215/Active site: Arg #status predicted  
 F;294-318, 334-353, 377-394, 465-478, 504-521/Disulfide bonds: #status experimental  
 F;336, 349, 360/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status experimental  
 F;548, 569/Binding site: N-acetylgalactosamine (Asp, Asn) #status experimental

Query Match 97.9%; Score 951; DB 1; Length 576;  
 Best Local Similarity 94.9%; Pred. No. 2.7e-79;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IPPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFILV 51  
 Db 36 IPPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINORFILV 95

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 111  
 Db 96 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 155

QY 112 GNYDRLEQLAGNLRNLELNGPLERASISALYYSTGTGTLPTLARSFFIICQMISEAAR 171  
 Db 156 GNYDRLEQLAGNLRNLELNGPLERASISALYYSTGTGTLPTLARSFFIICQMISEAAR 215

QY 172 FOYIEGMRTRIRYNRRS 189  
 Db 216 FOYIEGMRTRIRYNRRS 233

RESULT 2  
 RLCGSAG  
 agglutinin precursor - castor bean  
 N;Contents: rRNA N-glycosidase (EC 3.2.2.22)  
 C;Species: Ricinus communis (castor bean)  
 C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C;Accession: A24261; A24210  
 R;Robertus, L.M.; Lamb, I.F.; Pappin, D.J.C.; Lord, J.M.  
 J. Biol. Chem. 260, 15682-15686, 1985  
 A;Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.  
 A;Reference number: A24261; PMID:86059449; PMID:2999130  
 A;Accession: A24261  
 A;Molecule type: mRNA  
 A;Residues: 1-564 <ROB>  
 A;Cross-references: GB:M12089; NID:gl169700; PIDN:AAA33869.1; PID:gl169701

R;Araki, T.; Yoshioka, Y.; Funatsu, G.  
 Biochim. Biophys. Acta 872, 277-285, 1986  
 A;Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.  
 A;Reference number: A24210  
 A;Accession: A24210  
 A;Molecule type: protein  
 A;Residues: 303-325, 'P', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 553-572, 'E', 529-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F' <FUN>  
 C;Comment: This protein has strong agglutinating activity and weak cytotoxicity compared with ricin.  
 C;Superfamily: ricin; rRNA N-glycosidase homology  
 C;Keywords: duplication; glycoprotein; hydrolase; lectin; RNA binding; seed  
 F;1-24/Domain: signal sequence #status predicted <SIG>  
 F;25-290/Product: agglutinin chain A #status predicted <ACH>  
 F;35-281/Domain: rRNA N-glycosidase homology <RNG>  
 F;303-564/Product: agglutinin chain B #status experimental <BCH>  
 F;319-361, 362-402, 405-443, 450-485, 489-528, 531-564/Region: 40-residue repeats  
 F;34, 259/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;104, 147, 231, 232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
 F;200, 203/Active site: Glu, Arg #status predicted  
 F;282-306, 322-341, 365-382, 453-466, 492-509/Disulfide bonds: #status predicted  
 F;324, 337, 348/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status predicted  
 F;397, 437/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;536, 557/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 88.7%; Score 861.5; DB 1; Length 564;  
 Best Local Similarity 86.9%; Pred. No. 4.2e-71;  
 Matches 172; Conservative 7; Mismatches 8; Indels 11; Gaps 2;

QY 2 IPPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFILV 51  
 Db 25 IPPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINORFILV 84

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 111  
 Db 85 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRYYTFAFG 144

QY 112 GNYDRLEQLAGNLRNLELNGPLERASISALYYSTGTGTLPTLARSFFIICQMISEAAR 171  
 Db 145 GNYDRLEQLAGNLRNLELNGPLERASISALYYSTGTGTLPTLARSFFIICQMISEAAR 203

QY 172 FOYIEGMRTRIRYNRRS 189  
 Db 204 FOYIEGMRTRIRYNRRS 221

RESULT 3  
 RLZST  
 rRNA N-glycosidase (EC 3.2.2.22) alpha-trichosanthin precursor [validated] - Mongolian snake  
 N;Alternate names: alpha-TCS; type I ribosome-inactivating protein  
 C;Species: Trichosanthin kirilowii (Mongolian snake-gourd)  
 C;Date: 30-Sep-1988 #sequence\_revision 26-Jan-1996 #text\_change 23-Mar-2001  
 C;Accession: JT0566; A36274; JCI093; A36273; JT0003  
 R;Shaw, P.C.; Yung, M.H.; Zhu, R.H.; Ho, W.K.K.; Ng, T.B.; Yeung, H.W.  
 Gene 97, 267-272, 1991  
 A;Title: Cloning of trichosanthin cDNA and its expression in Escherichia coli.  
 A;Reference number: JT0566; PMID:91153657; PMID:1999291  
 A;Accession: JT0566  
 A;Molecule type: mRNA  
 A;Residues: 1-289 <SHA>  
 A;Cross-references: GB:M34858; NID:gl70536; PIDN:AAA34207.1; PID:gl70537  
 A;Experimental source: tuber  
 R;Chow, T.P.; Feldman, R.A.; Lovett, M.; Piatak, M.  
 J. Biol. Chem. 265, 8670-8674, 1990  
 A;Title: Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I ribosome-inactivating protein.  
 A;Reference number: A36274; PMID:90256790; PMID:2341400  
 A;Accession: A36274  
 A;Molecule type: DNA  
 A;Residues: 1-233, 'T', 235-246, 'M', 248-289 <CHO>  
 A;Cross-references: GB:J05434; NID:gl70534; PIDN:AAA34206.1; PID:gl70535  
 R;Zheng, H.G.; Wang, B.; Shao, P.Z.; Yang, X.R.  
 Acta Genet. Sin. 21, 42-51, 1994  
 A;Title: Cloning and DNA sequencing of the gene encoding Trichosanthin.  
 A;Reference number: JCI093; PMID:94271613; PMID:8003348  
 A;Accession: JCI093



Query Match	33.6%	Score 326;	DB 2;	Length 289;
Best Local Similarity	39.6%;	Pred. NO. 2.3e-22;		
Matches	72;	Conservative	44;	Mismatches 50;
				Indels 16;
				Gaps 5;



A>Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from  
A:Reference number: J70202  
A:Accession: J70202  
A:Molecule type: protein  
A:Residues: 1-201,203-251 <FUN>  
A>Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have  
R:Evans, G.; Mathiesen, A.; Sundan, A.  
J. Biol. Chem. 266, 6848-6852, 1991  
A>Title: Direct molecular cloning and expression of two distinct abrin A-chains.  
A:Reference number: A39761; MUID:91201329; PMID:2016300  
A:Accession: A39761  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 'E',2-251 <EVE>  
A:Cross-references: GB:54872  
A>Note: residues 1-8 were derived from the synthesized primer  
R:Kimura, M.; Sumizawa, T.; Funatsu, G.  
Biosci. Biotechnol. Biochem. 57, 166-169, 1993  
A>Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic  
A:Reference number: JCI398; MUID:93169023; PMID:7763422  
A:Contents: seeds  
A:Accession: JCI398  
A:Molecule type: protein  
A:Residues: 261-347,'T',349-351,'A',353-357,'L',359-528 <KIM>  
A:Experimental source: seed  
R:Evans, G.; Mathiesen, A.; Sundan, A.  
submitted to the EMBL Data Library, October 1990  
A>Description: Direct molecular cloning of two distinct abrin A-chains.  
A:Reference number: S14471  
A:Accession: S14472  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 'ME',2-251 <EV2>  
A:Cross-references: EMBL:X54873; NID:gl5090; PIDN:CAA38655.1; PID:gl6091  
R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.  
FEBS Lett. 309, 115-118, 1992  
A>Title: The complete primary structure of abrin-a B chain.  
A:Reference number: S24133; MUID:92371656; PMID:1505674  
A:Accession: S24133  
A:Molecule type: protein  
A:Residues: 262-297,'Y',299-426,'L',428-466,'P',468-482,'L',484-528 <CHE>  
R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.  
Eur. J. Biochem. 240, 564-569, 1996  
A>Title: Probing the domain structure of abrin-a by tryptic digestion.  
A:Reference number: S74110; MUID:97008945; PMID:8856055  
A:Accession: S74110  
A:Molecule type: protein  
A:Residues: 89-108;154-172 <LIN>  
A:Experimental source: seed  
A:Accession: S74111  
A:Molecule type: protein  
A:Residues: 262-276,'X',278-280;329-348;369-388;399-418 <LIW>  
A:Experimental source: seed  
C:Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which inhibits  
training receptors on the cell surface. The A and B chains are linked by a single disulfide  
C:Superfamily: ricin; rRNA N-glycosidase homology  
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic acid  
F:1-251/Product: abrin-a chain A #status experimental <ACH>  
F:7-245/Domain: rRNA N-glycosidase homology <RNG>  
F:261-528/Product: abrin-a chain B #status experimental <BCH>  
F:283-325;326-366;369-407;414-449;453-492;495-528/Region: 40-residue repeats  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted  
F:164,167/Active site: Glu, Arg #status predicted  
F:247-269;286-305;329-346;417-430;456-473/Disulfide bonds: #status predicted  
F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted  
F:361,401/Binding site: carboxylate (Asn) (covalent) #status experimental  
F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match  
Best Local Similarity 31.3%; Score 303.5; DB 1; Length 528;  
Matches 76; Conservative 26; Mismatches 64; Indels 19; Gaps 5;

QY 10 INFETAGATVQSYTNFIRAVGRLL-----TVLPNRVGLPINQRFILVLSNHAELSV 61  
DB 5 IKFSTEGATQSYKQFIETALRERLGGLIHDPVLPDPTTLOERNRYITVLSNSDTSI 64  
QY 62 TLALDVTNAYVVGVRAGNSAYFFH--PDNQDAEAIHLFTDVQNRVYFAFGNVDRLAQ 119  
DB 65 EVGIDVTNAYVAVRAGTQSYFLRDAFSSASD-----YLFGTG-DQSLPFGYGDLE 118  
QY 120 LAGNLRNIEIGNGLEBAISALYYSTGGTQLPTLARSFFICIMISEARFQVIEGEM 179  
DB 119 WAHQSQOIPLEGLOALTHGIS---FFRSGGNDNEKARTLIVIQMAEAAARFRYISNRV 175  
QY 180 RTRIR 184  
DB 176 RVSQ 180  
RESULT 12  
S23519  
beta-luffin - smooth loofah  
C:Species: Luffa cylindrica (smooth loofah)  
C>Date: 17-Apr-1993 #sequence revision 17-Apr-1993 #text\_change 20-Aug-1999  
C:Accession: S23519; S23113  
R:Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koizumi, A.  
Plant Mol. Biol. 19, 887-889, 1992  
A>Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-inactivating  
A:Reference number: S23519; MUID:92353400; PMID:1643290  
A:Accession: S23519  
A:Molecule type: mRNA  
A:Residues: 1-278 <KAT>  
A:Cross-references: EMBL:X62372; NID:gl9149; PIDN:CAA44230.1; PID:gl9150  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:26-264/Domain: rRNA N-glycosidase homology <RNG>  
Query Match 30.9%; Score 300.5; DB 2; Length 278;  
Best Local Similarity 35.3%; Pred. No. 4.7e-20;  
Matches 66; Conservative 43; Mismatches 57; Indels 21; Gaps 4;  
QY 10 INFETAGATVQSYTNFIRAVGRLL-----TVLPNRVGLPINQRFILVLSNHAEL 59  
DB 24 VSFSLGADSKSYKFIETALRERLGGLIHDPVLPDPTTLOERNRYITVLSNSDTSI 80  
QY 60 SVTLALDVTNAYVVGVRAGNSAYFFH--PDNQDAEAIHLFTDVQNRVYFAFGNVDRLAQ 119  
DB 81 AITMAIDVTNAYVVGVRAGNSAYFFH--PDNQDAEAIHLFTDVQNRVYFAFGNVDRLAQ 137  
QY 120 LAGNLRNIEIGNGLEBAISALYYSTGGTQLPTLARSFFICIMISEARFQVIEGEM 179  
DB 138 AAGKVRKPIGLGFRAFDSAITSLPHYDS-----TAAAGAFIVIIQTABASRKFVIEGOI 192  
QY 180 RTRIR 186  
DB 193 IERIPKN 199  
RESULT 13  
JN0108  
luffin-b - smooth loofah  
C:Species: Luffa cylindrica (smooth loofah)  
C>Date: 04-Sep-1998 #sequence revision 04-Sep-1998 #text\_change 07-May-1999  
C:Accession: JN0108  
R:Islam, M.R.; Hirayama, H.; Funatsu, G.  
Agric. Biol. Chem. 55, 229-238, 1991  
A>Title: Complete amino acid sequence of luffin-b, a ribosome-inactivating protein from  
A:Reference number: JN0108; MUID:91248488; PMID:1368666  
A:Accession: JN0108  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-250 <ISL>  
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology  
F:5-246/Domain: rRNA N-glycosidase homology <RNG>  
Query Match 30.2%; Score 293.5; DB 2; Length 250;







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:55 ; Search time 6.30467 Seconds  
(without alignments)  
1409.756 Million cell updates/sec

Title: US-10-083-336A-6  
Perfect score: 971  
Sequence: 1 MIFPKQYPIINFITAGATVQ.....ARFQYIEGEMRTIRYNRRS 189

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	951	97.9	576	1 RICI_RICCO	P02879 ricinus com
2	861.5	88.7	564	1 AGGL_RICCO	P06750 ricinus com
3	336	34.6	289	1 RPTC_TRIKI	P09989 trichosanthe
4	327.5	33.7	562	1 ABRC_ABRPR	P28590 abrus prece
5	326	33.6	289	1 RIPS_TRIKI	P24478 trichosanthe
6	323	33.3	527	1 ABRB_ABRPR	P06077 abrus prece
7	320.5	33.0	282	1 RIP2_BRVDI	P98184 bryonia dio
8	307.5	31.7	563	1 NIGB_SAMNI	P33183 sambucus ni
9	305	31.4	290	1 RIP1_BRVDI	P33185 bryonia dio
10	303.5	31.3	528	1 ABRA_ABRPR	P11140 abrus prece
11	293.5	30.2	250	1 RIPB_LUCFY	P22851 luffa cylin
12	281.5	29.0	286	1 RIP1_CUCFI	P36077 abrus prece
13	274.5	28.3	277	1 RIPA_LUCFY	P00465 luffa cylin
14	273	28.1	286	1 RIP1_MONCH	P16094 momordica c
15	272	28.0	285	1 RIP2_MOMBA	P29339 momordica b
16	264	27.2	254	1 MLA_VISAL	P81446 viscum albu
17	257	26.5	294	1 RIP1_TRIAN	P56626 trichosanthe
18	243.5	25.1	316	1 RIPG_GELMU	P33186 gelonium mu
19	186.5	19.2	294	1 RIPA_PHYAM	Q03464 phytoacca
20	174	17.9	278	1 RIPP_MIRUA	P21326 mirabilis j
21	170	17.5	313	1 RIP1_PHYAM	P10297 phytoacca
22	168.5	17.4	261	1 RIPS_PHYAM	P23339 phytoacca
23	131	13.5	253	1 RIP7_SAPOF	Q41391 saponaria o
24	127	13.1	253	1 RIP5_SAPOF	Q41389 saponaria o
25	125	12.9	310	1 RIP2_PHYAM	Q40772 phytoacca
26	124	12.8	299	1 RIP6_PHYAM	P20656 saponaria o
27	121	12.5	292	1 RIP2_SAPOF	P27559 saponaria o
28	116.5	12.0	293	1 RIP0_DIAAC	P24476 dianthus ca
29	114	11.7	280	1 RIP2_HORVU	P04399 hordeum vul
30	110	11.3	236	1 RIP3_SAPOF	P27560 saponaria o
31	108	11.1	319	1 SLTA_BP933	P09385 bacterioph
32	107	11.0	280	1 RIP1_HORVU	P22244 hordeum vul
33	103.5	10.7	300	1 RIP3_MAIZE	P25891 zea mays (m

## RESULT 1

RICI\_RICCO STANDARD; PRT; 576 AA.  
AC P02879; P02880;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase)  
(EC 3.2.2.22); Ricin B chain].  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86067214; PubMed=2999712;  
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,  
RA Weaver R.F.;  
RT "Genomic cloning and characterization of a ricin gene from Ricinus  
communis";  
RL Nucleic Acids Res. 13:8019-8033(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92163016; PubMed=1371405;  
RA Tregear J.W., Roberts L.M.;  
RT "The lectin gene family of Ricinus communis: cloning of a functional  
ricin gene and three lectin pseudogenes";  
RL Plant Mol. Biol. 19:515-525(1992).  
RN [3]  
RP SEQUENCE OF 12-576 FROM N.A.  
RX MEDLINE=85179479; PubMed=3838723;  
RA Lamb A., Roberts L.M., Lord J.M.;  
RT "Nucleotide sequence of cloned cDNA coding for preproricin";  
RL Eur. J. Biochem. 148:265-270(1985).  
RN [4]  
RP SEQUENCE OF 36-302.  
RA Yoshitake S., Funatsu G., Funatsu M.;  
RT "Isolation and sequences of peptic peptides, and the complete  
sequence of ile chain of ricin-D";  
RL Agric. Biol. Chem. 42:1267-1274(1978).  
RN [5]  
RP SEQUENCE OF 315-576.  
RA Funatsu G., Kimura M., Funatsu M.;  
RT "Primary structure of Ala chain of ricin D";  
RL Agric. Biol. Chem. 43:2221-2224(1979).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.  
RX MEDLINE=90344223; PubMed=1368517;  
RA Kimura Y., Kusuoku H., Tada M., Takagi S., Funatsu G.;  
RT "Structural analyses of sugar chains from ricin A-chain variant";  
RL Agric. Biol. Chem. 54:157-162(1990).  
RN [7]  
RP REVIEW.  
RX MEDLINE=21480122; PubMed=11595634;  
RA Olsnes S., Kozlov J.V.;

## ALIGNMENTS

RT "Ricin."; CC  
 RL Toxicon 39:1723-1728 (2001). CC  
 RN [8] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS). CC  
 RX MEDLINE=87165983; PubMed=3558397; CC  
 RA Monfort W., Villafranca J.E., Monzingo A.F., Ernst S.R., Katzin B., CC  
 RA Rutember E., Xuong N.H., Hamlin R., Robertus J.D.; CC  
 RL "The three-dimensional structure of ricin at 2.8 A."; CC  
 J. Biol. Chem. 262:5398-5403 (1987). CC  
 [9] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC  
 RX MEDLINE=91352004; PubMed=1881891; CC  
 RA Katzin B.J., Collins E.J., Robertus J.D.; CC  
 RL "Structure of ricin A-chain at 2.5 A."; CC  
 Proteins 10:251-259 (1991). CC  
 [10] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN. CC  
 RX MEDLINE=91352005; PubMed=1881882; CC  
 RA Rutember E., Robertus J.D.; CC  
 RL "Structure of ricin B-chain at 2.5-A resolution."; CC  
 Proteins 10:260-269 (1991). CC  
 [11] CC  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN. CC  
 RX MEDLINE=95082010; PubMed=7990130; CC  
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J., CC  
 RA Paupcit R.A.; CC  
 RL "X-ray structure of recombinant ricin A-chain at 1.8-A resolution."; CC  
 J. Mol. Biol. 244:410-422 (1994). CC  
 [12] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215. CC  
 RX MEDLINE=96374222; PubMed=8780513; CC  
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M., CC  
 RA Molina-Svinth M.C., Robertus J.D.; CC  
 RL "Structure and activity of an active site substitution of ricin A CC  
 chain."; CC  
 Biochemistry 35:11098-11103 (1996). CC  
 [13] CC  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC  
 RX MEDLINE=97240820; PubMed=9086280; CC  
 RA Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Malne G.W., CC  
 RA Robertus J.D.; CC  
 RL "Structure-based identification of a ricin inhibitor."; CC  
 J. Mol. Biol. 266:1043-1049 (1997). CC  
 [14] CC  
 RP MUTAGENESIS. CC  
 RX MEDLINE=93165632; PubMed=1287657; CC  
 RA Kin Y., Robertus J.D.; CC  
 RL "Analysis of several key active site residues of ricin A chain by CC  
 mutagenesis and X-ray crystallography."; CC  
 Protein Eng. 5:775-779 (1992). CC  
 -!- FUNCTION: Ricin is highly toxic to animal cells and to a less CC  
 extent to plant cells. The A chain is responsible for inhibiting CC  
 protein synthesis through the catalytic inactivation of 60S CC  
 ribosomal subunits. It acts as a glycosylase that removes a CC  
 specific adenine residue from an exposed loop of 28S ribosomal CC  
 RNA. As this loop is involved in the binding of elongation CC  
 factors, the modified ribosomes are unable to support protein CC  
 synthesis. The A chain can inactivate a few thousand ribosomes CC  
 per minute, thus inactivating them faster than the cell can make CC  
 new ones. A single A-chain molecule can therefore kill an animal CC  
 cell. The B chain binds to cell receptors and facilitates the CC  
 entry into the cell of the A chain; B chains are also responsible CC  
 for cell agglutination (lectin activity). It binds to beta-D- CC  
 galactopyranoside moieties. CC  
 -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one CC  
 specific adenosine on the 28S rRNA. CC  
 -!- SUBUNIT: Disulfide-linked dimer of A and B chains. CC  
 CC -!- DOMAIN: The B chain is composed of two domains, each domain CC  
 consists of 3 homologous subdomains (alpha, beta, gamma). CC  
 CC -!- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE CC  
 MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271. CC  
 -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME- CC  
 INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY. CC

CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains. CC  
 CC -!- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS CC  
 WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3). CC  
 CC -!- DATABASE: NAME=Protein Spotlight; CC  
 CC NOTE=Issue 31 of February 2003; CC  
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt031.html". CC  
 CC ----- CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC  
 between the Swiss Institute of Bioinformatics and the EMBL outstation - CC  
 the European Bioinformatics Institute. There are no restrictions on its CC  
 use by non-profit institutions as long as its content is in no way CC  
 modified and this statement is not removed. Usage by and for commercial CC  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/ CC  
 or send an email to license@isb-sib.ch). CC  
 CC ----- CC  
 CC EMBL; X03179; CAA26939.1; - CC  
 CC EMBL; X52908; CAA37095.1; - CC  
 CC EMBL; X02388; CAA26230.1; - CC  
 CC EMBL; A12892; CAA01058.1; - CC  
 CC FIR; A24041; RLCSD. CC  
 CC PDB; 2AAI; 31-JAN-94. CC  
 CC PDB; 1AFG; 31-JAN-94. CC  
 CC PDB; 1FMP; 31-OCT-93. CC  
 CC PDB; 11FS; 14-JAN-98. CC  
 CC PDB; 11FT; 14-JAN-98. CC  
 CC PDB; 11FU; 14-JAN-98. CC  
 CC PDB; 1RTC; 31-OCT-93. CC  
 CC PDB; 1OBS; 16-JUN-97. CC  
 CC PDB; 1OBT; 16-JUN-97. CC  
 CC PDB; 1BR5; 02-SEP-98. CC  
 CC PDB; 1BR6; 02-SEP-98. CC  
 CC PDB; 1IL3; 16-JAN-02. CC  
 CC PDB; 1IL4; 16-JAN-02. CC  
 CC PDB; 1IL9; 16-JAN-02. CC  
 CC GlycoSuiteDB; P02879; - CC  
 CC InterPro; IPR000772; Ricin\_B\_lectin. CC  
 CC InterPro; IPR001574; RIP. CC  
 CC Pfam; PF00652; Ricin\_B\_lectin; 6. CC  
 CC Pfam; PF00161; RIP; 1. CC  
 CC PRINTS; PR00396; SHIGARICIN. CC  
 CC SMART; SM00458; RICIN; 2. CC  
 CC PROSITE; PS00231; RICIN\_B\_LECTIN; 2. CC  
 CC PROSITE; PS00275; SHIGA\_RICIN; 1. CC  
 CC Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; CC  
 KW Glycoprotein; Lectin; Signal; 3D-structure. CC  
 FT SIGNAL 1 35 CC  
 FT CHAIN 36 302 CC  
 FT PEPTIDE 303 314 CC  
 FT CHAIN 315 576 CC  
 FT CHAIN 321 448 CC  
 FT DOMAIN 451 575 CC  
 FT REPEAT 331 373 CC  
 FT REPEAT 374 414 CC  
 FT REPEAT 417 449 CC  
 FT REPEAT 462 497 CC  
 FT REPEAT 501 540 CC  
 FT REPEAT 543 570 CC  
 FT ACT\_SITE 212 212 CC  
 FT DISULFID 294 318 CC  
 FT DISULFID 334 353 CC  
 FT DISULFID 377 394 CC  
 FT DISULFID 465 478 CC  
 FT DISULFID 504 521 CC  
 FT CARBOHYD 45 45 CC  
 FT CARBOHYD 271 271 CC  
 FT CARBOHYD 409 409 CC  
 FT CARBOHYD 449 449 CC  
 FT CARBOHYD 76 76 CC  
 FT CONFLICT 551 551 CC  
 FT STRAND 43 47 CC  
 FT TURN 49 50 CC

CC ----- CC  
 CC N-LINKED (GLCNAC. . .). CC  
 CC /FTIG-CAR 000080. CC  
 CC N-LINKED (GLCNAC. . .) (IN MINOR FORM). CC  
 CC /FTIG-CAR 000081. CC  
 CC N-LINKED (GLCNAC. . .). CC  
 CC N-LINKED (GLCNAC. . .). CC  
 CC E -> D (IN REF. 3). CC  
 CC A -> R (IN REF. 3). CC

Query Match 97.9%; Score 951; DB 1; Length 576;  
 Best Local Similarity 94.9%; Pred. No. 3.2e-80;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFATTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 51  
 DB 36 IFPKQYPIINFATTAGATVQSYNFIKAVRGLTGTGADVRHEIPVLPNRVGLPINQRFILV 95

QY 52 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 111  
 DB 96 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 155

QY 112 GNYDRLEQLAGNLRNIELNGPLERALSALYYSTGGTOLPTLARSFICIMISEAR 171  
 DB 156 GNYDRLEQLAGNLRNIELNGPLERALSALYYSTGGTOLPTLARSFICIMISEAR 215

QY 172 FOYIEGEMTRIRYNRS 189  
 DB 216 FOYIEGEMTRIRYNRS 233

RESULT 2  
 AGGL\_RICCO STANDARD; PRT; 564 AA.  
 AC P06750;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (SC 3.2.2.22); Agglutinin B chain].  
 OS Ricinus communis (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 OX NCBI\_TaxID=3988;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=86059449; PubMed=2999130;  
 RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;  
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";  
 RL J. Biol. Chem. 260:15682-15686(1985).  
 RN [2]  
 RN SEQUENCE OF 303-564.  
 RC TISSUE=Seed;  
 RA Araki T., Yoshioka Y., Funatsu G.;  
 RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";  
 RL Biochim. Biophys. Acta 872:277-285(1986).  
 RN [3]  
 RN SEQUENCE OF 303-337.  
 RX MEDLINE=80178723; PubMed=6768555;  
 RA Lin T.-S., Li S.-L.;  
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";  
 RL Eur. J. Biochem. 105:453-459(1980).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; M2089; AAA33869.1; -;  
 CC EMBL; S40368; AAB22584.1; -;

DR PIR; A24261; RLCSAG.  
 DR HSP; P02879; LBR6.  
 DR GlycoSitedB; P06750; -;  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 KW Glycoprotein; Lectin; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 290 AGGLUTININ A CHAIN.  
 FT PROPEP 291 302 LINKER PEPTIDE.  
 FT CHAIN 303 564 AGGLUTININ B CHAIN.  
 FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 319 361 1-ALPHA.  
 FT REPEAT 362 402 1-BETA.  
 FT REPEAT 405 437 1-GAMMA.  
 FT REPEAT 450 485 2-ALPHA.  
 FT REPEAT 489 528 2-BETA.  
 FT REPEAT 531 558 2-GAMMA.  
 FT ACT\_SITE 200 200 BY SIMILARITY.  
 FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 322 341 BY SIMILARITY.  
 FT DISULFID 365 382 BY SIMILARITY.  
 FT DISULFID 453 466 BY SIMILARITY.  
 FT DISULFID 492 509 BY SIMILARITY.  
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .).  
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. .).  
 FT CONFLICT 331 331 F -> T (IN REF. 2).  
 FT CONFLICT 362 362 N -> D (IN REF. 2).  
 FT CONFLICT 374 374 R -> G (IN REF. 2).  
 FT CONFLICT 404 404 R -> T (IN REF. 2).  
 FT CONFLICT 552 552 F -> V (IN REF. 2).  
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match 88.7%; Score 861.5; DB 1; Length 564;  
 Best Local Similarity 86.9%; Pred. No. 5.6e-72;  
 Matches 172; Conservative 7; Mismatches 8; Indels 11; Gaps 2;

QY 2 IFPKQYPIINFATTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 51  
 DB 25 IFPKQYPIINFATTAGATVQSYNFIKAVRGLTGTGADVRHEIPVLPNRVGLPINQRFILV 84

QY 52 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 111  
 DB 85 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 144

QY 112 GNYDRLEQLAGNLRNIELNGPLERALSALYYSTGGTOLPTLARSFICIMISEAR 171  
 DB 145 GNYDRLEQLAGNLRNIELNGPLERALSALYYSTGGTOLPTLARSFICIMISEAR 203

QY 172 FOYIEGEMTRIRYNRS 189  
 DB 204 FOYIEGEMTRIRYNRS 221

RESULT 3  
 RIPT\_TRIKI STANDARD; PRT; 289 AA.  
 ID RIPT TRIKI  
 AC P09989;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ribosome-inactivating protein alpha-trichosanthin precursor (rRNA N-glycosidase) (BC 3.2.2.22) (Alpha-TCS).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales, Cucurbitaceae; Trichosanthes.  
 ON NCBI\_TaxID=3677;  
 RX SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz;  
 RX MEDLINE=91153657; PubMed=199291;  
 RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;  
 RT "Cloning of trichosanthin cDNA and its expression in *Escherichia*  
 RL *coli*."; Gene 97:267-272(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Maximowicz; TISSUE=Leaf;  
 RX MEDLINE=90256790; PubMed=2341400;  
 RA Chow T., Feldman R.A., Lovett M., Piatak M.;  
 RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a  
 RL type I ribosome-inactivating protein."; J. Biol. Chem. 265:8670-8674(1990).  
 RN [3]  
 RP SEQUENCE OF 24-270.  
 RC STRAIN=Maximowicz; TISSUE=Tuberos root;  
 RX MEDLINE=90256789; PubMed=2341399;  
 RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,  
 RT Wu P., Hwang K., Piatak M.;  
 RL "Primary amino acid sequence of alpha-trichosanthin and molecular  
 models for abrin A-chain and alpha-trichosanthin."; J. Biol. Chem. 265:8665-8669(1990).  
 RN [4]  
 RP SEQUENCE OF 24-270.  
 RC TISSUE=Tuberos root;  
 RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,  
 RT Tian G.Y., Ni C.Z.;  
 RL "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and  
 application."; Pure Appl. Chem. 58:789-798(1986).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).  
 RX MEDLINE=94344957; PubMed=8066085;  
 RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;  
 RT "Structure of trichosanthin at 1.88-A resolution."; Proteins 19:4-13(1994).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).  
 RX MEDLINE=95344383; PubMed=7619070;  
 RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;  
 RT "Studies on crystal structures, active-centre geometry and  
 depurinatin mechanism of two ribosome-inactivating proteins."; Biochem. J. 309:285-298(1995).  
 RL -!- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS  
 CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT  
 INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M34858; AAA34207.1; -;  
 DR EMBL; J05434; AAA34206.1; -;  
 DR PIR; J0566; RLTT.  
 DR PDB; 1MRJ; 07-FEB-95.  
 DR PDB; 1MRK; 07-FEB-95.  
 DR PDB; 1TCS; 10-JUL-95.  
 DR PDB; 1U4G; 28-JAN-03.

DR PDB; 1WLI; 21-JAN-03.  
 DR PDB; 1QD2; 24-APR-00.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PRO0396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 270  
 FT PROPEP 271 289  
 FT ACT\_SITE 183  
 FT CONFLICT 57 60  
 FT CONFLICT 82 84  
 FT CONFLICT 87 87  
 FT CONFLICT 92 92  
 FT CONFLICT 143 144  
 FT CONFLICT 196 196  
 FT CONFLICT 215 216  
 FT CONFLICT 231 231  
 FT CONFLICT 234 234  
 FT CONFLICT 246 266  
 FT CONFLICT 247 247  
 FT STRAND 25 28  
 FT TURN 30 31  
 FT HELIX 34 46  
 FT TURN 47 47  
 FT STRAND 50 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 89 92  
 FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT TURN 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT TURN 142 142  
 FT HELIX 145 147  
 FT STRAND 150 150  
 FT HELIX 152 163  
 FT TURN 164 165  
 FT HELIX 167 180  
 FT TURN 181 181  
 FT HELIX 182 186  
 FT STRAND 187 187  
 FT HELIX 188 195  
 FT TURN 196 196  
 FT STRAND 202 202  
 FT HELIX 206 226  
 FT TURN 227 230  
 FT STRAND 231 239  
 FT TURN 241 242  
 FT STRAND 245 250  
 FT TURN 251 252  
 FT HELIX 254 258  
 FT TURN 259 259  
 FT STRAND 260 260  
 FT TURN 263 263  
 FT TURN 266 268  
 FT TURN 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;  
 SQ SEQUENCE

Query Match 34.6%; Score 336; DB 1; Length 289;  
 Best Local Similarity 38.9%; Pred. No. 7.9e-24;  
 Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVRGLTVLPN-----RVGLPINQRFILVELSNAEL 59  
 Db 25 VSPRLSGATSSSYGVFISNLR---KALPNERKLYDIPLLRSSLPQSGQRYALIHLYNVADE 81  
 QY 60 SVTLALDVTNAYVGVYRAGNSAYFFHPDNCEDA-EALTHLFTVQVRYTFAGGNYDRLE 118  
 Db 82 TTSVADVTNVIYMGVRAGTSTVF---NEASATEAKYVFKDMARKVTLPLYSNGVIERLQ 138  
 QY 119 QLAGNLRNIELNGPLLEAISALYYSTGCTQPLTLARSFICIQMISEARFQVIEGE 178  
 Db 139 TAAGKIRENIPGLPALDSAITLIFYNAN-----SAASALMWLIQSTSEAAKYKIEQQ 193  
 QY 179 MRTRI 183  
 Db 194 IGRKV 198

RESULT 4  
 ABRC ABRPR STANDARD; PRT; 562 AA.  
 AC P28590;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abrin-c precursor [contains: Abrin-c A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-c B chain].  
 OS Abrus precatorius (Indian licorice) (Crab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=91266957; PubMed=2050149;  
 RX Wood K.A., Lord J.M., Wawrzynczak E.J., Piatok M.;  
 RA "Preproabrin: genomic cloning, characterisation and the expression of  
 RT the A-chain in *Escherichia coli*.";  
 RL Eur. J. Biochem. 198:723-732(1991).  
 CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 CC BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; X55667; CAA39202.1; --  
 DR PIR; S16022; S16022.  
 DR HSSP; P11140; 1ABR.  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR SMART; SM00458; RICIN; 2.  
 DR PROSITE; P850231; RICIN B LECTIN; 2.  
 DR PROSITE; P850275; SHIGA-RICIN; 1.  
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;

KW Glycoprotein; Lectin; Signal; Pyroglutamate carboxylic acid.  
 FT SIGNAL 1 34 BY SIMILARITY.  
 FT CHAIN 35 285 ABRIN C A CHAIN (BY SIMILARITY).  
 FT PEPTIDE 286 295 LINKER PEPTIDE (BY SIMILARITY).  
 FT CHAIN 296 562 ABRIN C B CHAIN (BY SIMILARITY).  
 FT DOMAIN 307 434 RICIN B-TYPE LECTIN 1.  
 FT DOMAIN 437 561 RICIN B-TYPE LECTIN 2.  
 FT REPEAT 317 359 1-ALPHA.  
 FT REPEAT 360 400 1-BETA.  
 FT REPEAT 403 435 1-GAMMA.  
 FT REPEAT 448 483 2-ALPHA.  
 FT REPEAT 487 526 2-BETA.  
 FT REPEAT 529 562 2-GAMMA.  
 FT ACT\_SITE 198 198 BY SIMILARITY.  
 FT DISULFID 281 303 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 320 339 BY SIMILARITY.  
 FT DISULFID 363 380 BY SIMILARITY.  
 FT DISULFID 451 464 BY SIMILARITY.  
 FT DISULFID 490 507 BY SIMILARITY.  
 FT MOD\_RES 35 35 PYROGLUTAMATE CARBOXYLIC ACID  
 FT CARBOHYD 234 234 (BY SIMILARITY).  
 FT CARBOHYD 395 395 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 435 435 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;

Query Match 33.7%; Score 327.5; DB 1; Length 562;  
 Best Local Similarity 43.9%; Pred. No. 1.1e-32;  
 Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;

QY 6 QYPLINFTTAGATVQSYTNFIRAVRGLT-----VLENRVGLPINQRFILVELSNAH 57  
 Db 35 QDQVKEFTGATSSQYKQFIEALRQLTGLIHDPVLPDPTTVEERNYITVELSENSE 94  
 QY 58 ELSVTALDVTNAYVGVYRAGNSAYFFHPDNCQDAEAIHFLTDTVQVRYTFAGGNYDR 117  
 Db 95 RESIEVGIDVTNAYVAVYRAGSQSYFL---RDAPASASTYLPFGTQ-RYSLRFDGSGYDL 150  
 QY 118 EQLAGNLRNIELNGPLLEAISALYYSTGCTQPLTLARSFICIQMISEARFQVIEG 177  
 Db 151 ERWAHQTRFEEISLGLQALTHAIS---FLRGSANDEKARTLIVLIQWASEAARYISN 207  
 QY 178 EMRTIR 184  
 Db 208 RVGSIR 214

RESULT 5  
 RIPS TRIKI STANDARD; PRT; 289 AA.  
 AC P24478;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein karasurin precursor (rRNA  
 DE N-glycosidase) (EC 3.2.2.22).  
 OS Trichosanthes kirilowii (Mongolian snake-gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.  
 OX NCBI\_TaxID=3677;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Root tuber;  
 RX MEDLINE=97356562; PubMed=9212998;  
 RX Mizukami H., Iida K., Kondo T., Ogihara Y.;  
 RT "Cloning and bacterial expression of a gene encoding ribosome-  
 RT inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes  
 RT kirilowii var. japonica";  
 RL Biol. Pharm. Bull. 20:711-713 (1997).  
 RN [2]  
 RP SEQUENCE OF 24-270.  
 RX MEDLINE=92005921; PubMed=1914000;

MEDLINE=93132798; PubMed=8421313;  
Hung C.-H., Lee M.-C., Lee T.-C.; Lin J.-Y.;  
RT "Primary structure of three distinct isoabrinins determined by cDNA  
sequencing. Conservation and significance.";  
J. Mol. Biol. 229:263-267(1993).  
[2]  
SEQUENCE OF 260-527.  
RP  
RN TISSUE=Seed;  
RC  
RX MEDLINE=93169023; PubMed=7763422;  
RA Kimura M., Sumizawa T., Funatsu G.;  
RT "The complete amino acid sequences of the B-chains of abrin-a and  
RT abrin-b, toxic proteins from the seeds of Abrus precatorius.";  
RL Biosci. Biotechnol. Biochem. 57:166-169(1993).  
CC  
CC !- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
CC  
CC ABRIN-A IS MORE TOXIC THAN RICIN.  
CC  
CC !- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
PRECEDES ENDOCYTOSIS.  
CC  
CC !- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
specific adenosine on the 28S rRNA.  
CC  
CC !- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
CC  
CC !- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
CC  
CC !- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
CC  
CC !- SIMILARITY: Contains 2 ricin B-type lectin domains.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; M98345; AAA32625.1; -;  
PIR; S32430; S32430.  
DR HSP; P11140; IABR.  
DR  
DR InterPro: IPR000772; Ricin\_B\_lectin.  
DR InterPro: IPR001574; RIP.  
DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS0231; RICIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
KW Glycoprotein; Lectin; Pyrrolidone carboxylic acid.  
FT CHAIN 1 250 ABRIN-B A CHAIN.  
FT PEPTIDE 251 260 LINKER PEPTIDE.  
FT CHAIN 261 527 ABRIN-B B CHAIN.  
FT DOMAIN 272 399 RICIN B-TYPE LECTIN 1.  
FT DOMAIN 402 526 RICIN B-TYPE LECTIN 2.  
FT REPEAT 282 324 1-ALPHA.  
FT REPEAT 325 365 1-BETA.  
FT REPEAT 368 400 1-GAMMA.  
FT REPEAT 413 448 2-ALPHA.  
FT REPEAT 452 491 2-BETA.  
FT REPEAT 494 527 2-GAMMA.  
FT ACT\_SITE 163 163 BY SIMILARITY.  
FT DISULFID 246 268 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 285 304 BY SIMILARITY.  
FT DISULFID 328 345 BY SIMILARITY.  
FT DISULFID 416 429 BY SIMILARITY.  
FT DISULFID 455 472 BY SIMILARITY.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY  
SIMILARITY).  
FT CARBOHYD 110 110 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 360 360 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 400 400 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 282 282 N -> D (IN REF. 2).

FT CONFLICT 291 291 D -> N (IN REF. 2).  
 FT CONFLICT 350 351 AE -> PQ (IN REF. 2).  
 FT CONFLICT 378 378 S -> N (IN REF. 2).  
 FT CONFLICT 426 426 L -> M (IN REF. 2).  
 FT CONFLICT 428 428 Y -> D (IN REF. 2).  
 FT CONFLICT 431 431 N -> S (IN REF. 2).  
 FT CONFLICT 431 431 N -> S (IN REF. 2).  
 FT CONFLICT 484 484 R -> K (IN REF. 2).  
 FT CONFLICT 484 484 R -> K (IN REF. 2).  
 FT CONFLICT 491 491 N -> S (IN REF. 2).  
 FT CONFLICT 493 493 H -> Y (IN REF. 2).  
 FT CONFLICT 502 502 R -> Q (IN REF. 2).  
 FT CONFLICT 509 509 E -> Q (IN REF. 2).  
 FT CONFLICT 513 513 H -> W (IN REF. 2).  
 FT CONFLICT 516 516 H -> T (IN REF. 2).  
 SQ SEQUENCE 527 AA; 59114 MW; 3253AE490C99494A CRC64;

Query Match 33.3%; Score 323; DB 1; Length 527;  
 Best Local Similarity 43.9%; Pred. No. 2.6e-22;  
 Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;

QY 6 QYPIINFTAGATVQSYTNFIRAVRGRLT-----VLNRYGLPINOQFVLVLSNHA 57  
 Db 1 QDQVIFTEGATGSSQYKIEALRQLTGLHIGIPVLPDPTTLQERNRYIISVELSNSD 60  
 QY 58 ELSVLTALDVNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFATGNYDRL 117  
 Db 61 TESIEAGIDVSNAYVAYRAGNSYFL---RDAPTSASRYLFTGTQ-QYSLRFGNSYIDL 116  
 QY 118 EQLAGNIRENIELGNGPLEBAISALYYSTGGTQPLTARSFICQIMISEAARFOYIEG 177  
 Db 117 ERLAROTROQIPLGLQALRAHSFPL-----QSGTDDQEIARTLIVIQMASEAARYRFSY 172  
 QY 178 EMRTIRYN 186  
 Db 173 RVGSIRTN 181

## RESULT 7

RIP2\_BRYDI  
 ID RIP2\_BRYDI STANDARD; PRT; 282 AA.  
 AC P98184; Q9S8J0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (SD2).  
 OS Bryonia dioica (Red Bryonia).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 OX NCBI\_TaxID=3652;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Siegal C.B., Gawlak S.L., Marquardt H.;  
 RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant  
 RT Bryonia dioica."  
 RL Patent number US5597569, 28-JAN-1997.  
 RN [2]  
 RP SEQUENCE OF 22-42.  
 RC TISSUE=Root;  
 RX MEDLINE=95151812; PubMed=7849072;  
 RA Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
 RA Marquardt H.;  
 RT "Characterization of ribosome-inactivating proteins isolated from  
 RT Bryonia dioica and their utility as carcinoma-reactive  
 RT immunoconjugates."  
 RL Bioconj. Chem. 5:423-429(1994).  
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----NOT ANNOTATED\_CDS-----  
 DR EMBL; I34238; -; NOT ANNOTATED\_CDS.  
 DR HSSP; P09989; 1MRJ.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP.1.  
 DR PRINTS; P00396; SHIGARICIN.  
 DR PROSITE; PS00295; SHIGA\_RICIN; 1.  
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
 KW Multigene family; Glycoprotein; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 282 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.  
 FT ACT SITE 183 183 BY SIMILARITY.  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;

Query Match 33.0%; Score 320.5; DB 1; Length 282;  
 Best Local Similarity 46.0%; Pred. No. 2.1e-22;  
 Matches 81; Conservative 24; Mismatches 54; Indels 17; Gaps 7;

QY 10 INETTAGATVQSYTNFIRAVRGRLTV-LPNRVGLPINQ-----RFLVLSNHAELSV 61  
 Db 24 INFSLGATGATYKTFIRNLRLTKLTGTPRVVDIPVLRNAAAGLARFQLVLTNYGESV 83  
 QY 62 TLALDVNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFATGNYDRLVQLA 121  
 Db 84 TVALDVNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRVTFATGNYDRLVQLA 139  
 QY 122 GNL-RENIELGNGPLEBAISALYYSTGGTQPLTARSFICQIMISEAARFOYIE 176  
 Db 140 GRISRENIELGNGPLEBAISALYYSTGGTQPLTARSFICQIMISEAARFOYIE 191

## RESULT 8

NIGB\_SAMNI  
 ID NIGB\_SAMNI STANDARD; PRT; 563 AA.  
 AC P33183; P33184; P91542;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Nigrin b precursor (Agglutinin V) (SNAV) [Contains: Nigrin b A chain  
 DE (rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain].  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Barb;  
 RC MEDLINE=96215449; PubMed=8647092;  
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;  
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V  
 RT (nigrin b), a GalNAC-specific type-2 ribosome-inactivating protein  
 RT from the bark of elderberry (Sambucus nigra).";  
 RL Eur. J. Biochem. 237:505-513(1996).  
 RN [2]  
 RP SEQUENCE OF 26-49 AND 298-321.  
 RC TISSUE=Barb;  
 RX MEDLINE=94003077; PubMed=8400135;  
 RA Gibbs T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R.,  
 RA Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;  
 RT "Isolation and partial characterization of nigrin b, a non-toxic  
 RT novel type 2 ribosome-inactivating protein from the bark of Sambucus  
 RT nigra L.";  
 RL Plant Mol. Biol. 22:1181-1186(1993).  
 CC -1- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN



CC PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN  
 CC SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE  
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE  
 CC BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES  
 CC ENOCYTOSIS.  
 CC  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-  
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.  
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U41299; AAB39475.1; -;  
 CC PIR; S37382; S37382.  
 CC InterPro; IPR000772; Ricin\_B\_lectin.  
 CC InterPro; IPR001574; RIP.  
 CC Pfam; PF00652; Ricin\_B\_lectin; 6.  
 CC Pfam; PF00161; RIP; 1.  
 CC PRINTS; PR00396; SHIGARICIN.  
 CC SMART; SM00458; RICIN; 2.  
 CC PROSITE; PS00275; SHIGA\_RICIN; 1.  
 CC PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;  
 CC Glycoprotein; Lectin; Signal.  
 CC SIGNAL 1 25  
 CC CHAIN 26 297 NIGRIN B A CHAIN.  
 CC CHAIN 298 563 NIGRIN B B CHAIN.  
 CC DOMAIN 305 431 RICIN B-TYPE LECTIN 1.  
 CC DOMAIN 434 559 RICIN B-TYPE LECTIN 2.  
 CC REPEAT 316 356 1-ALPHA.  
 CC REPEAT 357 397 1-BETA.  
 CC REPEAT 400 432 1-GAMMA.  
 CC REPEAT 445 482 2-ALPHA.  
 CC REPEAT 486 524 2-BETA.  
 CC REPEAT 527 554 2-GAMMA.  
 CC ACT\_SITE 188 188 INTERCHAIN (BY SIMILARITY).  
 CC DISULFID 274 302 INTERCHAIN (BY SIMILARITY).  
 CC DISULFID 319 338 BY SIMILARITY.  
 CC DISULFID 360 377 BY SIMILARITY.  
 CC DISULFID 448 463 BY SIMILARITY.  
 CC DISULFID 489 506 BY SIMILARITY.  
 CC CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 39 39 K -> V (IN REF. 2).  
 CC SEQUENCE 563 AA; 62300 MW; F250CB24621BF14 CRG64;  
 CC  
 CC Query Match 31.7%; Score 307.5; DB 1; Length 563;  
 CC Best Local Similarity 38.4%; Pred. No. 7.6e-21;  
 CC Matches 73; Conservative 36; Mismatches 56; Indels 25; Gaps 5;  
 CC  
 CC QY 7 YPIINTTACATVQSYNTIRAVR-----GRLTVLPKRVGLPQINQRIIVLSNH 56  
 CC DB 28 YPVSFNLGKASATYQDFLSNRKATVATGTVEVNGLPVLRRESEYQVKSRFVLVLTNY 87  
 CC QY 57 AELSVTLALDVNTAYVGVYVGRAGNSAYFFHPDNOEDAEAI--THLFDVQNRVTFARQNY 114  
 CC DB 88 NGNTVTLAVDVNTLYVAVSGNANSIFF-----KDTEVQKSNLFGVTQKN-TLSTGNY 141  
 CC QY 115 DRLEQLAGNLRNTEIENGNGLEPAISALYYSTGGTQLPFLTARSFICIQMISEAARFOY 174

DB 142 DNLETAANTRRSELELGPSPUDGAITSLYHGD-----SVARSLLVVIQMVSEARFRY 194  
 QY 175 IEGEMTRIR 184  
 DB 195 IEQEVRSLSQ 204  
 CC  
 CC RESULT 9  
 CC RIPI\_BRYDI  
 CC ID RIPI\_BRYDI STANDARD; PRT; 290 AA.  
 CC AC P33185; O9S819;  
 CC DT 01-OCT-1993 (Rel. 27, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)  
 CC DE (EC 3.2.2.22) (BD1).  
 CC OS Bryonia dioica (Red bryony).  
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.  
 CC OX NCBI\_TaxID=3652;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
 CC RC TISSUE=Leaf;  
 CC EX MEDLINE=97228081; PubMed=9115985;  
 CC RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,  
 CC RA Siegall C.B.;  
 CC RT "Molecular, biological, and preliminary structural analysis of  
 CC RT recombinant bryodin I, a ribosome-inactivating protein from the plant  
 CC RT Bryonia dioica.";  
 CC RL Biochemistry 36:3095-3103 (1997).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RA Siegall C.B.;  
 CC RT "Cloning and expression of a gene encoding bryodin I from Bryonia  
 CC RT dioica.";  
 CC RL Patent number US5541110, 30-JUL-1996.  
 CC RN [3]  
 CC RP SEQUENCE OF 24-66.  
 CC RC TISSUE=Seed;  
 CC RX MEDLINE=89326691; PubMed=2753596;  
 CC RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,  
 CC RA Lappi D.;  
 CC RT "N-terminal sequence of some ribosome-inactivating proteins.";  
 CC RL Int. J. Pept. Protein Res. 33:263-267 (1989).  
 CC RN [4]  
 CC RP SEQUENCE OF 24-43.  
 CC RC TISSUE=Root;  
 CC RX MEDLINE=95151812; PubMed=7849072;  
 CC RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,  
 CC RA Marquardt H.;  
 CC RT "Characterization of ribosome-inactivating proteins isolated from  
 CC RT Bryonia dioica and their utility as carcinoma-reactive  
 CC RT immunoconjugates.";  
 CC RL Bioconj. Chem. 5:423-429 (1994).  
 CC CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS  
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS.  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO  
 CC PRODUCE A SHORTER PROTEIN.  
 CC -!- BIOTECHNOLOGY: Especially useful as immunotoxin for  
 CC pharmacological applications as it has low toxicity in rats and  
 CC mice but is potent once inside target cells.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC

or send an email to license@isb-sib.ch).

-----NOT ANNOTATED CDS.

CC CC EMBL; I24020; -; NOT ANNOTATED CDS.  
 DR PIR; S16491; S16491.  
 DR PDB; 1BRV; 04-MAR-98.  
 DR InterPro; IPR001574; RIP.  
 DR Pfam; PF00161; RIP; 1.  
 DR PRINTS; PRO0396; SHIGARICIN.  
 DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
 KW plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;  
 KW 3D-structure; Multigene family; Glycoprotein; Signal.  
 FT SIGNAL 1 23  
 FT CHAIN 24 270 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.  
 FT PROPEP 271 290 MISSING IN MATURE PROTEIN.  
 FT ACT\_SITE 183 183 BY SIMILARITY.  
 FT ACT\_SITE 212 212  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MUTAGEN 212 212 E-X; REDUCES ACTIVITY 10-FOLD.  
 FT CONFLICT 61 65 RSSIS -> LRHXI (IN REF. 3).  
 FT STRAND 25 28  
 FT TURN 30 31  
 FT HELIX 34 46  
 FT TURN 47 47  
 FT STRAND 50 54  
 FT TURN 55 56  
 FT STRAND 57 60  
 FT HELIX 66 69  
 FT STRAND 70 76  
 FT TURN 78 79  
 FT STRAND 82 88  
 FT TURN 89 92  
 FT STRAND 93 99  
 FT TURN 100 101  
 FT STRAND 102 105  
 FT HELIX 109 114  
 FT TURN 115 117  
 FT TURN 120 121  
 FT STRAND 124 127  
 FT HELIX 134 141  
 FT TURN 142 142  
 FT HELIX 145 147  
 FT STRAND 150 150  
 FT HELIX 152 163  
 FT TURN 164 165  
 FT STRAND 167 186  
 FT HELIX 187 187  
 FT STRAND 188 196  
 FT STRAND 202 202  
 FT HELIX 206 213  
 FT TURN 214 214  
 FT TURN 215 225  
 FT STRAND 226 230  
 FT STRAND 231 239  
 FT TURN 241 242  
 FT STRAND 245 250  
 FT TURN 251 252  
 FT HELIX 254 257  
 FT TURN 258 259  
 FT STRAND 260 260  
 FT STRAND 263 263  
 FT HELIX 266 268  
 SQ SEQUENCE 290 AA; 31788 MW; B966CD9C031A42DB CRC64;  
 Query Match 31.4%; Score 305; DB 1; Length 290;  
 Best Local Similarity 35.7%; Pred. No. 5.8e-21;  
 Matches 65; Conservative 45; Mismatches 56; Indels 16; Gaps 4;  
 Qy 10 INFTTAGATQSYNFTFRAVGRGRTVLPNVRGLPI-----NORFILVELSNHAEISVT 62  
 Db 25 VSFRLSGATTTSYGVIKKNLREALPYERKVNYPILLRSSISGSGRYTLHLTLNVADETIS 84  
 Qy 63 LALDVTNAYVVGVRAGNSAYFFHFDNQEDA-EAITHLFTDVQNEYTFAGFGNYDRLEQLA 121

Db 85 VADVTVNVMGVLADVSFF---NEASATEAAKVFKDAAKKVTLPLPSYNGYERLQTAA 141  
 Qy 122 GNLRNIELNGEPLERAIISALYYSTGGTQPLTARSFFICIMISEAARFOVIEGEMRT 181  
 Db 142 GKIRENIFGLPALDSAITLTYTAS-----SAASALLVLQSTAESARYKEIEQIGK 196  
 Qy 182 RI 183  
 Db 197 RV 198  
 RESULT 10  
 ABRA ABRPR STANDARD; PRT; 528 AA.  
 ID AC P11140; P28589;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-a B chain].  
 OS Abrus precatorius (Indian licorice) (Grab's eye).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
 OX NCBI\_TaxID=3816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93132798; PubMed=8421313;  
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;  
 RT "Primary structure of three distinct isoforms determined by cDNA sequencing. Conservation and significance.";  
 RL J. Mol. Biol. 229:263-267(1993).  
 RN [2]  
 RP SEQUENCE OF 1-251.  
 RC TISSUE=Seed;  
 RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;  
 RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from the seeds of Abrus precatorius.";  
 RL Agric. Biol. Chem. 52:1095-1097(1988).  
 RN [3]  
 RP SEQUENCE OF 1-251 FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE=91201329; PubMed=2016300;  
 RA Eysen G., Mathiesen A., Sundan A.;  
 RT "Direct molecular cloning and expression of two distinct abrin A-chains.";  
 RL J. Biol. Chem. 266:6848-6852(1991).  
 RN [4]  
 RP SEQUENCE OF 262-528.  
 RX MEDLINE=92371656; PubMed=1505674;  
 RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;  
 RT "The complete primary structure of abrin-a B chain.";  
 RL FEBS Lett. 309:115-118(1992).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).  
 RX MEDLINE=95333188; PubMed=7608980;  
 RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;  
 RT "Crystal structure of abrin-a at 2.14 A.";  
 RL J. Mol. Biol. 250:354-367(1995).  
 CC -I- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA.  
 CC ABRIN-A IS MORE TOXIC THAN RICIN.  
 CC -I- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.  
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.  
 CC -I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
 CC -I- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-

Query Match 31.3%; Score 303.5; DB 1; Length 528;  
Best Local Similarity 41.1%; Pred. No. 1.6e-20;  
Matches 76; Conservative 26; Mismatches 64; Indels 19

QY 10 INFTAGATVQSYTNFIRAVRGRL-----TVLPNRVGLPINQRFILVELSNHAEISV 61  
 Db 5 IKPSTEGATSQSYKQFIEALREELRGGLIHDIPVDPDPTLQERNRYITVELSNSDTEI 64  
 QY 62 TLALDVTNAYVVGVRAGNSAYFFH--PDNCEDEAETHLFTDVQNRYYTFAFGNVDRLQ 119  
 Db 65 EVGIDVTNAYVVRAGTQSYFLRDPSSASD-----YLFGTG-DQHSLEFFGYTYGDLER 118  
 QY 120 LAGNLENIENGPLEEALISALYYSTGTQPTLARSFIIQIMISEAARFOYIEGEM 179  
 Db 119 WAHQSRQIPLGLQALTHGIS---FFRSGNDNEEKARTLIVIQWVAARFRIYSNRV 175  
 QY 180 RTRIR 184  
 Db 176 RVSIIQ 180

## RESULT 11

ID RIPB\_LUCFY STANDARD; PRT; 250 AA.  
 AC P22851;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ribosome-inactivating protein luffin-B (rRNA N-glycosidase)  
 DE (EC 3.2.2.22)  
 OS Luffa cylindrica (Smooth loofah) (Sponge gourd).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.  
 OX NCBI\_TaxID=3670;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Seed;  
 RX MEDLINE=91248488; PubMed=1368666;  
 RA Islam M.R., Hirayama H., Funatsu G.;  
 RT "Complete amino acid sequence of luffin-b, a ribosome-inactivating  
 RT protein from sponge gourd (Luffa cylindrica) seeds.";  
 RL Agric. Biol. Chem. 55:229-238 (1991).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC PIR: JN0108; JN0108.  
 DR HSSP: P16094; 1AHC.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: P000396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin.  
 FT ACT\_SITE 160 160 BY SIMILARITY.  
 SQ SEQUENCE 250 AA; 27293 MW; F01A8DC9A1078700 CRC64;

Query Match 30.2%; Score 293.5; DB 1; Length 250;  
 Best Local Similarity 33.7%; Pred. No. 5.6e-20;  
 Matches 63; Conservative 46; Mismatches 57; Indels 21; Gaps 4;

QY 10 INFTAGATVQSYTNFIRAVRGRL-----TVLPNRVGLPINQRFILVELSNHAEISV 59  
 Db 3 VPSLGSADSKSYKFTALRKALPSKEVSNIPLLPSASGA---SRYLMLQSNYDAK 59  
 QY 60 SVTLALDVTNAYVVGVRAGNSAYFFHFDNCEDEAETHLFTDVQNRYYTFAFGNVDRLQ 119  
 Db 60 AITWALDVTNVIYMGVLVNSTSYF---ANESDAKLASQYVFKGSTLTIPYSNGYERLQN 116  
 QY 120 LAGNLENIENGPLEEALISALYYSTGTQPTLARSFIIQIMISEAARFOYIEGEM 179  
 Db 117 AAKIKREKIPGLGFALDSALTSIFHYDS-----TAAAAAEVLQTTAEASRKYIEGQI 171  
 QY 180 RTRIRYN 186  
 Db 172 IERIPKN 178

## RESULT 12

ID RIPL\_CUCFI STANDARD; PRT; 286 AA.  
 AC Q9FRX4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)  
 DE (EC 3.2.2.22)  
 OS Cucumis ficaria.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OX NCBI\_TaxID=131071;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada T., Ohki S.T., Osaki T.;  
 RT "Cloning and analysis of a cDNA coding a putative ribosome-  
 RT inactivating protein from Cucumis ficaria.";  
 RL Plant Biotechnol. 17:337-340 (2000).  
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
 CC specific adenosine on the 28S rRNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 CC TYPE 1 RIP SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR HSSP: AB045560; BAB19677.1; -  
 DR EMSP: P16094; 1AHC.  
 DR InterPro: IPR001574; RIP.  
 DR Pfam: PF00161; RIP; 1.  
 DR PRINTS: P000396; SHIGARICIN.  
 DR PROSITE: PS00275; SHIGA\_RICIN; 1.  
 KW plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;  
 KW Toxin; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 286 PUTATIVE RIBOSOME-INACTIVATING PROTEIN.  
 FT ACT\_SITE 185 185 BY SIMILARITY.  
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 286 AA; 31771 MW; 4EFD4966E604DA41 CRC64;

Query Match 29.0%; Score 281.5; DB 1; Length 286;  
 Best Local Similarity 32.4%; Pred. No. 8.4e-19;  
 Matches 60; Conservative 46; Mismatches 64; Indels 15; Gaps 3;

QY 10 INFTAGATVQSYTNFIRAVRGRLTVLPNRVGLPI-----NORFILVELSNHAEISV 62  
 Db 28 VPSLGSADSKSYKFTSMRNALPNAGDIYNIPLVPSISGSRYYLMQLSNYEGRNTIT 87  
 QY 63 LALDVTNAYVVGVRAGNSAYFFHFDNCEDEAETHLFTDVQNRYYTFAFGNVDRLQ 122  
 Db 88 MAVDVTNVIYMGVLVNGTSYFF---NETDAQLASKFVFGTKSITLTPYSGNYKQLQSVAR 144  
 QY 123 NURENIEGNGLPGLAEALISALYYSTGTQPTLARSFIIQIMISEAARFOYIEGEMRTR 182  
 Db 145 KERDSIPLGFALDSALTSIYYDSRSAPI-----AFLVIQTAAEARYKYIEKQIIDR 199  
 QY 183 IRYNR 187  
 Db 200 ISVSK 204

## RESULT 13

```

RIPA_LURCY
ID RIPA_LURCY STANDARD; PRT; 277 AA.
AC Q00465;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein luffin-alpha precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22)
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX NCBI_TaxID=3670;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=92288316; PubMed=1600156;
RA Kataoka J., Habuka N., Miyano M., Masuta C., Koizumi A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
RT inactivating protein from luffa cylindrica.";
RL Plant Mol. Biol. 18:1199-1202(1992).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62371; CAA44229.1; -
CC PIR; S22494; S22494.
CC HSP; P16094; LAHC.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PRO0396; SHIGARICIN.
CC PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
FT Toxin; Signal.
FT SIGNAL 1 19
FT CHAIN 20 277 RIBOSOME-INACTIVATING PROTEIN LUFFIN-
FT ALPHA.
FT ACT_SITE 179 179 BY SIMILARITY.
FT SEQUENCE 277 AA; 30212 MW; EA17FC27998C25AC CRC64;
SQ
Query Match 28.3%; Score 274.5; DB 1; Length 277;
Best Local Similarity 33.3%; Pred. No. 3.6e-18;
Matches 63; Conservative 42; Mismatches 61; Indels 23; Gaps 5;
QY 10 INFTTAGATVQSYTNFIRAVRGRLTVLPNRVGLPIN-----QRFILVELSNHAE 58
DB 22 VRFSLGSSSTVSYSKFIGDLR---KALPSN-GPVNITLLSSAGASRYTLWLSNYDG 77
QY 59 LSVTLALDVTNAYVGYVAGNAGSYFHPDQDAEALITLFTDQVQRYTFAGGNYDRLE 118
DB 78 KAITVAVDVTNVMGYLVNYSYFF---NESDAKLASQYVFGKSTIVLPYSGNYEKLU 134
QY 119 QLAGNLRNEMTELGNGPLEBAISALYYVYTGGLTGLPLARSFICIMISEAARFOYIEGE 178
DB 135 TAAGKIREKIPGLFPALDSAITLTFHYS-----TAAAAAFVLIOTTAEGRFKYIEQ 189
QY 179 MTRIRYNR 187
DB 190 IIRISKNQ 198

```

RESULT 14  
RIP1\_MOMCH

```

AC P16094; P24697;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein momordin I precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=91159486; PubMed=2001404;
RA Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
RT protein.";
RL Biochim. Biophys. Acta 1088:311-314(1991).
RN [2]
RP SEQUENCE OF 24-38.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins.";
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [3]
RP SEQUENCE OF 24-70.
RC TISSUE=Seed;
RX MEDLINE=89005108; PubMed=3262509;
RA Casellas P., Dussossa D., Palasca A.I., Barbieri L.,
RA Guilleminot J.C., Ferrara P., Bolognesi A., Cennini P., Stirpe F.;
RT "Trichokirin, a ribosome-inactivating protein from the seeds of
RT Trichosanthes kirilowii Maximowicz. Purification, partial
RT characterization and use for preparation of immunotoxins.";
RL Eur. J. Biochem. 176:581-588(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=94356447; PubMed=8075985;
RA Ren J., Wang Y., Dong Y., Stuart D.I.;
RT "The N-glycosidase mechanism of ribosome-inactivating proteins
RT implied by crystal structures of alpha-momorcharin.";
RL Structure 2:7-16(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
RX MEDLINE=94192822; PubMed=8143869;
RA Husain J., Tickle I.J., Wood S.P.;
RT "Crystal structure of momordin, a type I ribosome inactivating
RT protein from the seeds of Momordica charantia.";
RL FEBS Lett. 342:154-158(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=95344383; PubMed=7619070;
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RT depurinating mechanism of two ribosome-inactivating proteins.";
RL Biochem. J. 309:285-298(1995).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X57682; CAA40869.1; -
CC PIR; S14273; RLPUGG.

```

DR PDB; 1AHA; 22-JUN-94.  
DR PDB; 1AHB; 22-JUN-94.  
DR PDB; 1AHC; 22-JUN-94.  
DR PDB; 1AMC; 22-JUN-94.  
DR PDB; 1MOM; 31-MAY-94.  
DR PDB; 1MRG; 07-FEB-95.  
DR PDB; 1MRH; 07-FEB-95.  
DR PDB; 1MRI; 07-FEB-95.  
DR PDB; 1F8Q; 21-NOV-01.  
DR GlycositedB; P16094; --  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
KW Glycoprotein; 3D-structure.  
FT SIGNAL 1 23  
FT CHAIN 24 269  
FT PROPEP 270 286  
FT ACT SITE 183 183  
FT CARBOHYD 250 250  
FT STRAND 25 28  
FT HELIX 34 47  
FT STRAND 50 54  
FT TURN 55 56  
FT STRAND 57 60  
FT HELIX 66 69  
FT STRAND 70 76  
FT TURN 78 79  
FT STRAND 82 88  
FT TURN 89 92  
FT STRAND 93 99  
FT TURN 100 101  
FT STRAND 102 105  
FT HELIX 109 114  
FT TURN 115 117  
FT TURN 120 121  
FT STRAND 124 127  
FT HELIX 134 141  
FT STRAND 145 147  
FT STRAND 150 150  
FT HELIX 152 162  
FT TURN 163 163  
FT HELIX 167 186  
FT STRAND 187 187  
FT HELIX 188 195  
FT TURN 196 197  
FT STRAND 202 202  
FT HELIX 206 225  
FT TURN 226 230  
FT STRAND 231 238  
FT TURN 240 241  
FT STRAND 246 250  
FT TURN 251 252  
FT HELIX 254 258  
FT TURN 259 259  
FT STRAND 260 260  
FT STRAND 263 263  
FT HELIX 266 268  
SQ SEQUENCE 286 AA; 31532 MW; E1B013ABEBC216CF CRC64;  
Query Match 28.1%; Score 273; DB 1; Length 286;  
Best Local Similarity 34.8%; Pred. No. 5.1e-16;  
Matches 63; Conservative 40; Mismatches 62; Indels 16; Gaps 4;  
QY 10 INFTAGATQVSQYTNFIRAVRGRLTVLPNRVGLPI-----NORFILVELSHAEISVT 62  
DB 25 VSRFGADPRSYGMFVKDLGNALPFREKVVNIPLLPVSUAGRYLLMHLFNVYDGKTI 84  
QY 63 LALDVNAYVGVYRAGNSAYFFHPDQEDAE-AATHLFTDVQNYRTAFGNGYDRLEQLA 121  
DB 85 VAVDTNVTNYINGYLADTTSYFF--NEPAELASQYVFRDARRKTIPLPSGNYERLQIAA 141

QY 122 GMLRENIELNGPLLEBAISALYYSTGGTQLPTLARSFIICTQMISEAARFOYIEGEMRT 181  
DB 142 GKPREKIPIGIPALDSASTLLHYDS-----TAAAGALLVLIQTAAARFKYIEQIQIE 196  
QY 182 R 182  
DB 197 R 197  
RESULT 15  
RIP2 MOMB  
ID RIP2 MOMB STANDARD; PRT; 286 AA.  
AC P29339;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Ribosome-inactivating protein momordin II precursor (rRNA  
DE N-glycosidase) (EC 3.2.2.22).  
OS Momordica balsamina (Bitter melon) (Balsam pear).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Cucurbitales; Cucurbitaceae; Momordica.  
OX NCBI\_TaxID=3672;  
RN [1]  
SEQUENCE FROM N.A.  
RC TISSUE=Seed;  
RX MEDLINE=93027170; PubMed=1408771;  
RA Ortigao M., Better M.;  
RT "Momordin II, a ribosome inactivating protein from Momordica  
RL Nucleic Acids Res. 20:4662-4662(1992).  
CC -! CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S rRNA.  
CC -! SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC TYPE 1 RIP SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; Z12175; CAA78166.1; --  
DR PIR; S25560; S25560.  
DR PDB; 1CF5; 07-JUN-99.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
KW Plant defense; protein synthesis inhibitor; Hydrolase; Toxin; Signal;  
KW 3D-structure.  
FT SIGNAL 1 23  
FT CHAIN 24 286  
FT RIBOSOME-INACTIVATING PROTEIN MOMORDIN  
FT II.  
FT ACT SITE 181 181  
FT BY SIMILARITY.  
SQ SEQUENCE 286 AA; 32031 MW; 3B89FFIAE6B25986 CRC64;  
Query Match 28.0%; Score 272; DB 1; Length 286;  
Best Local Similarity 34.2%; Pred. No. 6.3e-18;  
Matches 63; Conservative 39; Mismatches 66; Indels 16; Gaps 4;  
QY 10 INFTAGATQVSQYTNFIRAVRGRLTVLPNRVGLPI-----NORFILVELSHAEISVT 62  
DB 25 VNFDELSTATAKTYKTEIDPRATLPFSHKVYDIPLYSTISDSRRFLDLTSYAVETIS 84  
QY 63 LALDVNAYVGVYRAGNSAYFFHPDQEDAEATHLFTDVQNYRTAFGNGYDRLEQLAG 122  
DB 85 VALDVNTNAYVAVRTDVSYFF---KESPEANVILFKGTR-KITLPYTGNYENLQTAH 140  
QY 123 NLRENIELNGPLLEBAISALYYSTGGTQLPTLARSFIICTQMISEAARFOYIEGEMRT 182

Db 141 KIRENID/LGLPALSSAITTLFYNA-----QSAPSALLVLIQTTAEARFKYIERHVAKY 195

Qy 183 IRYN 186

Db :  
196 VATN 199

Search completed: February 10, 2004, 16:23:24  
Job time : 7.30467 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 25.2187 Seconds  
(without alignments)  
1933.961 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MTFPKQYPIINTTAGATVQ.....ARFQYIEGEMTRIRNRRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SPREMBL 23:\*

- 1: sp\_archea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951	97.9	541	10	Q41174
2	374.5	38.6	580	10	Q94BW3
3	372.5	38.4	580	10	Q94BW4
4	370.5	38.2	581	10	Q94BW5
5	368.5	38.0	549	10	Q95V22
6	334	34.4	289	10	Q94KE4
7	332	34.2	289	10	Q41216
8	329.5	33.9	563	10	Q04367
9	326	33.6	247	10	Q9LRE3
10	323.5	33.3	564	10	Q9AVR2
11	318.5	32.8	528	10	Q06076
12	317	32.6	270	10	Q8LPV7
13	310.5	32.0	252	10	Q38760
14	307.5	31.7	563	10	Q945S2
15	307.5	31.7	563	10	Q8GT32
16	305	31.4	270	10	Q41611

17	303.5	31.3	252	10	Q38761
18	302	31.1	565	10	Q04071
19	300.5	30.9	278	10	Q00980
20	298.5	30.7	547	10	Q9M6E9
21	297.5	30.6	251	10	Q96236
22	296.5	30.5	251	10	Q96237
23	293.5	30.2	566	10	Q04072
24	290.5	29.9	251	10	Q96235
25	273	28.1	264	10	Q9FSH2
26	273	28.1	570	10	Q41358
27	271	27.9	592	10	Q8W2E7
28	270	27.8	249	10	Q8LK05
29	270	27.8	286	10	Q9FUV7
30	267	27.5	570	10	Q22415
31	266.5	27.4	604	10	Q9M654
32	265	27.3	286	10	Q41257
33	263	27.1	565	10	Q8W243
34	260	26.8	254	10	Q8LK06
35	260	26.8	573	10	Q8W2E8
36	259	26.7	251	10	Q8LKQ4
37	259	26.7	293	10	Q8S452
38	259	26.7	569	10	Q93543
39	253	26.1	531	10	Q8RXH6
40	251	25.8	249	10	Q8RXH7
41	247	25.4	258	10	Q9S9E4
42	247	25.4	293	10	Q8VYU0
43	246.5	25.4	203	10	Q8RY69
44	246.5	25.4	275	10	Q8H1Y4
45	245.5	25.3	603	10	Q9M653

## ALIGNMENTS

RESULT 1

Q41174  
ID Q41174 PRELIMINARY; PRT; 541 AA.  
AC Q41174;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Proridin A chain (EC 3.2.2.22) (rRNA N-glycosidase)  
DE (Fragment).  
OS Ricinus communis (Castor bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
OX NCBI\_TaxID=3988;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92338377; PubMed=1633311;  
RA Roberts L.M., Tregear J.W., Lord J.M.;  
RT "Molecular cloning of ricin.";  
RL Targeted Diagn. Ther. 7:81-97(1992).  
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
CC EMBL; S40366; AAB22582.1; -.  
DR HSSP; P02879; IBR6.  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR InterPro; IPR001400; Somatotropin.  
DR Pfam; PF00852; Ricin\_B\_lectin; 6.  
DR Pfam; PF00161; RIP; 1.  
DR SMART; SM00458; RICIN; 2.  
DR PRINTS; PR00396; SHIGARICIN.  
DR PROSITE; PS00231; RICIN\_B\_LECTIN; 2.  
DR PROSITE; PS00275; SHIGA\_RICIN; 1.  
DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
KW Hydrolase; Toxin.  
FT NON TER 1  
SQ SEQUENCE 541 AA; 60281 MW; 2B7B2CDF1F2E9D9 CRC64;



Query Match 97.9%; Score 951; DB 10; Length 541;  
 Best Local Similarity 94.9%; Pred. No. 2.8e-82;  
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 51  
 DB 1 IFPKQYPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFG 111  
 DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFG 120

QY 112 GNYDELEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEAR 171  
 DB 121 GNYDELEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEAR 180

QY 172 FOYIEGEMTRIRYNRS 189  
 DB 181 FOYIEGEMTRIRYNRS 198

RESULT 2  
 Q94BW3  
 ID Q94BW3 PRELIMINARY; PRT; 580 AA.  
 AC Q94BW3  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin III precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AY039803; AAK82460.1; -  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR SMART; SM00458; RICIN; 2.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 KW Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match 38.6%; Score 374.5; DB 10; Length 580;  
 Best Local Similarity 48.1%; Pred. No. 3.7e-27;  
 Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56  
 DB 33 YQTVFTTKNATKTSYQFIEALRAQLASGEEPHGIPVNRERSTVPDSKRFLVLSNWA 92

QY 57 AELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 116  
 DB 93 ADSPVTLAVDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 149

QY 117 LEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEARFOVIE 176  
 DB 150 LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLLVVIQWVAEAVRFRFIE 207

Query Match 38.4%; Score 372.5; DB 10; Length 580;  
 Best Local Similarity 47.6%; Pred. No. 5.8e-27;  
 Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56  
 DB 33 YQTVFTTKNATKTSYQFIEALRAQLASGEEPHGIPVNRERSTVPDSKRFLVLSNWA 92

QY 57 AELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 116  
 DB 93 ADSPVTLAVDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 149

QY 117 LEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEARFOVIE 176  
 DB 150 LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLLVVIQWVAEAVRFRFIE 207

Query Match 38.4%; Score 372.5; DB 10; Length 580;  
 Best Local Similarity 47.6%; Pred. No. 5.8e-27;  
 Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56  
 DB 33 YQTVFTTKNATKTSYQFIEALRAQLASGEEPHGIPVNRERSTVPDSKRFLVLSNWA 92

QY 57 AELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 116  
 DB 93 ADSPVTLAVDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 149

QY 117 LEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEARFOVIE 176  
 DB 150 LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLLVVIQWVAEAVRFRFIE 207

QY 177 GEMETRI 183  
 DB 208 YRVRESI 214

RESULT 3  
 Q94BW4  
 ID Q94BW4 PRELIMINARY; PRT; 580 AA.  
 AC Q94BW4  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Type 2 ribosome-inactivating protein cinnamomin II precursor  
 DE (EC 3.2.2.22) (rRNA N-glycosidase).  
 OS Cinnamomum camphora (Camphor tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
 OX NCBI\_TaxID=13429;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yang Q., Gong Z.Z., Liu W.Y.;  
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
 RT genes encoding cinnamomin proteins and study of their expression  
 RT patterns";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
 CC SPECIFIC ADENOSINE ON THE 28S RNA.  
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
 DR EMBL; AY039802; AAK82459.1; -  
 DR InterPro; IPR000772; Ricin\_B\_lectin.  
 DR Pfam; PF00652; Ricin\_B\_lectin; 6.  
 DR SMART; SM00458; RICIN; 2.  
 DR PRINTS; PR00396; SHIGARICIN.  
 DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
 KW Hydroxylase; Signal; Toxin.  
 FT SIGNAL 1 32  
 FT CHAIN 33 580  
 FT SEQUENCE 580 AA; 64265 MW; 37E4289ECE0CBFF CRC64;

Query Match 38.4%; Score 372.5; DB 10; Length 580;  
 Best Local Similarity 47.6%; Pred. No. 5.8e-27;  
 Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56  
 DB 33 YQTVFTTKNATKTSYQFIEALRAQLASGEEPHGIPVNRERSTVPDSKRFLVLSNWA 92

QY 57 AELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 116  
 DB 93 ADSPVTLAVDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 149

QY 117 LEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEARFOVIE 176  
 DB 150 LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLLVVIQWVAEAVRFRFIE 207

Query Match 38.4%; Score 372.5; DB 10; Length 580;  
 Best Local Similarity 47.6%; Pred. No. 5.8e-27;  
 Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56  
 DB 33 YQTVFTTKNATKTSYQFIEALRAQLASGEEPHGIPVNRERSTVPDSKRFLVLSNWA 92

QY 57 AELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 116  
 DB 93 ADSPVTLAVDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 149

QY 117 LEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEARFOVIE 176  
 DB 150 LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLLVVIQWVAEAVRFRFIE 207

Query Match 38.4%; Score 372.5; DB 10; Length 580;  
 Best Local Similarity 47.6%; Pred. No. 5.8e-27;  
 Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY 7 YPIINFNTAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILVLSN-H 56  
 DB 33 YQTVFTTKNATKTSYQFIEALRAQLASGEEPHGIPVNRERSTVPDSKRFLVLSNWA 92

QY 57 AELSVTALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 116  
 DB 93 ADSPVTLAVDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYFAFGNVDYR 149

QY 117 LEQAGNLENIENGNGPLEEASALYYSTGTQPLTARSFIICMISEARFOVIE 176  
 DB 150 LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLLVVIQWVAEAVRFRFIE 207

OS Cinnamomum camphora (Camphor tree).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
NCBI\_TaxID=13429;  
[1]  
RN SEQUENCE FROM N.A.  
RA Yang Q., Gong Z.Z., Liu W.Y.;  
RP "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)  
RT genes encoding cinnamomin proteins and study of their expression  
RT patterns";  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
RC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AV039801; AAK82458.1; -;  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00652; Ricin\_B\_lectin; 5.  
DR Pfam; PF00161; RIP; 1.  
DR PRINTS; PR00396; SHIGARICIN.  
DR SMART; SM00458; RICIN; 2.  
DR PROSITE; PS50231; RICIN\_B\_LECTIN; 2.  
DR Hydrolyase; Signal; Toxin.  
FT SIGNAL 1 32  
FT CHAIN 33 581  
FT FT  
FT CHAIN 33 581  
SQ SEQUENCE 581 AA; 64215 MW; 5E8F5FB8FA3D196 CRC64;  
  
Query Match 38.2%; Score 370.5; DB 10; Length 581;  
Best Local Similarity 48.1%; Pred.No. 9e-27;  
Matches 90; Conservative 25; Mismatches 57; Indels 15; Gaps 5;  
  
QY 7 YPIINFTTAGATVQSNTFIRAVRGRLT-----VLPNRVGLPINQRFILVELSN-H 56  
Db 33 YQVFTTTRKATKTSYQTFEALRAQLASGEHPHGVMRERSTVPDSKRFILVELSNWA 92  
  
QY 57 AELSVTLLADVTNAYVYGYRAGNSAYFFHPDNOQDAEAIHLFTDVQNRXYTFAFGGNYDR 116  
Db 93 ADSFVTLAVDVTNAYVYRGTSGSQFPLEDNPD--PAIENLLPDTK-RYTFPFGSYTD 149  
  
QY 117 LEQIAGNLRENIELGNGPLREASIALYVYGGTGTLPFLARSFICICOMISEARFOYIE 176  
Db 150 LEGVAGRRREIILGMDPLENAISALMISNL--NQQRALARSILVVIQMAEAVRFRFIE 207  
  
QY 177 GEMRTRI 183  
Db 208 YRVGRSI 214  
  
RESULT 5  
Q9FV22 PRELIMINARY; PRT; 549 AA.  
ID AC Q9FV22;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Type II ribosome-inactivating protein cinnamomin (EC 3.2.2.22) (rRNA  
DE N-glycosidase) (Fragment).  
OS Cinnamomum camphora (Camphor tree).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.  
NCBI\_TaxID=13429;  
[1]  
RN SEQUENCE FROM N.A.  
RA Xie L., Liu W.-Y., Wang E.-D.;  
RP "Molecular cloning of cinnamomin A-, B-chain and the expression,  
RT purification, characterization and mutagenesis of the A-Chain";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RNA.  
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.  
DR EMBL; AF259548; AAF68978.2; -;  
DR HSPSP; P02879; 2AAL.

```

Db 82 TTSVADVTNVMYGRAGDTSYFF---NEASATEAAKYVFKDSMRKITLPYSGNVERLQ 138
QY 119 QLAGNREINELNGPLEFAISALYYSTGGTLPFLARSFICIMISEAARFOYIEGE 178
Db 139 TAAGKIRENIPGLPALDSAITFLFYNNAN-----SAASALMWLIQSTSEAAKYKFEQ 193
QY 179 MRTRI 183
Db 194 IGRV 198

RESULT 7
Q41216 PRELIMINARY; PRT; 289 AA.
AC Q41216;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE TRICHOSANTHIN (EC 3.2.2.22) (rRNA N-glycosidase).
GN TRICHOSANTHIN, TCS
OS Trichosanthin kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94271613; PubMed=8003348;
RA Zheng H., Wang B., Shaw P., Yeung H.;
RT "[Cloning and DNA sequencing of the gene encoding trichosanthin].";
RL I Chuan Hsueh Pao 21:42-51(1994).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; S70176; AAB31048.1; -.
DR HSSP; P09989; IMRU.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
SQ SEQUENCE 289 AA; 31650 MW; 286AC14D4B8CA175 CRC64;

Query Match 34.2%; Score 332; DB 10; Length 289;
Best Local Similarity 38.9%; Pred. No. 1.7e-23;
Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVRGRLTVLPN-----RVGLPINQRFILVELSNHAE 59
Db 25 VSFRLSGATSSSYGVFISNLK---KALPNRKLYDIPLLRSLPQSGRYALVHLNTYADE 81
QY 60 SVTLALDVTNNAVYVGRAGNSAYFFHPDNCEDA-EATHTLFTDVQNYTFAFGNVDRL 118
Db 82 TTSVADVTNVMYGRAGDTSYFF---NEASATEAAKYVFKDAMRKVTLPYSGNVERLQ 138
QY 119 QLAGNREINELNGPLEFAISALYYSTGGTLPFLARSFICIMISEAARFOYIEGE 178
Db 139 TAAGKIRENIPGLPALDSAITFLFYNNAN-----SAASALMWLIQSTSEAAKYKFEQ 193
QY 179 MRTRI 183
Db 194 IGRV 198

RESULT 8
O04367 PRELIMINARY; PRT; 563 AA.
AC O04367;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Ribosome inactivating protein precursor (EC 3.2.2.22) (rRNA
DE N-glycosidase).

```

```

OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.
OX NCBI_TaxID=4202;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98112023; PubMed=9450339;
RA Van Damme E.J., Roy S., Barre A., Rouge P., Van Leuven F.,
RA Peumans W.J.;
RT "The major elderberry (Sambucus nigra) fruit protein is a lectin
RT derived from a truncated type 2 ribosome-inactivating protein.";
RL Plant J. 12:1251-1260(1997).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; U76524; AAC15886.1; -.
DR HSSP; P02879; 2AA1.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP_lectin; 6.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 25
FT CHAIN 26 297
FT CHAIN 298 563
FT CHAIN 563 563
SQ SEQUENCE 563 AA; 62336 MW; 3ED2B6C08E796205 CRC64;

Query Match 33.9%; Score 329.5; DB 10; Length 563;
Best Local Similarity 40.5%; Pred. No. 7.2e-23;
Matches 77; Conservative 37; Mismatches 51; Indels 25; Gaps 5;

QY 7 YPINFTTAGATVQSYTNFIRAVR-----GRLTVLPNRVGLPINQRFILVELSNH 56
Db 28 YPSVSFNLAGAKSATYRDFLNRLRTIVATGYEVNGLPVLRRSEVQVKNREVLVLLNY 87
QY 57 AELSVTLALDVTNNAVYVGRAGNSAYFFHPDNCDEAI--THLFTDVQNYTFAFGNY 114
Db 88 NGNTVTTLAVDVTNLYVAFSANANSYFF-----KDATQLQKSNLVGTR-QHTLPFTGNY 141
QY 115 DELEQLAGNLEINELNGPLEFAISALYYSTGGTLPFLARSFICIMISEAARFOY 174
Db 142 DNLETAAGTRRSIELGSPLOGALTSLYDE-----SVARSLLVVIQMWSEARPRY 194
QY 175 IEGEMRTIR 184
Db 195 IEQEVRSRLQ 204

RESULT 9
Q9LRE3 PRELIMINARY; PRT; 247 AA.
AC Q9LRE3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Trichobakin (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
GN TBK.
OS Trichosanthes sp. Bac Kan 8-98.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=118182;
RN [1]
RP SEQUENCE FROM N.A.
RA Phan Van Chi, Nguyen Thuy Ha, Le Tran Binh;
RT "Genomic DNA Clone for mature typ-1 ribosome-inactivating protein from
RT Trichosanthes sp. sample 01 Bac Kan 8-98 Vien CNSH (Hanoi).";
RT Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.

```

[illegible]

OS *Abrus precatorius* (Indian licorice) (Crab's eye).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosid 1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.  
NCBI\_TaxID=3816;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=LEAF.  
RC MEDLINE=91201329; PubMed=2016300;  
RX Evensen G., Mathiesen A., Sundan A.;  
RA "Direct molecular cloning and expression of two distinct abrin A-  
RT chains".  
RT J. Biol. Chem. 266:6848-6852(1991).  
RL CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN  
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL  
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.  
CC ABRIN-A IS MORE TOXIC THAN RICIN.  
CC CC -!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT  
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT  
CC PRECEDES ENDOCYTOSIS.  
CC CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE  
CC SPECIFIC ADENOSINE ON THE 28S RNA.  
CC CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.  
CC CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN  
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).  
CC CC -!- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING  
CC PROTEINS. BELONGS TO TYPE 2 RIP.  
CC EMBL; X54872; CAA38654.1; -;  
CC HSSP; P11140; 1ABR.  
CC InterPro; IPR001574; RIP.  
CC Pfam; PF00161; RIP; 1.  
CC PROSITE; PR00396; SHIGARICIN.  
CC DR PROSITE; PS00275; SHIGA RICIN; 1.  
CC DR Hydroxylase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.  
CC KW CHAIN 1 252 ABRIN B, A CHAIN (BY SIMILARITY).  
CC FT NON\_TER 252 252  
CC SEQUENCE 252 AA; 28309 MW; BBFC846B9E92B5DE CRC64;  
Query Match 32.0%; Score 310.5; DB 10; Length 252;  
Best Local Similarity 43.2%; Pred. No. 1.6e-21;  
Matches 79; Conservative 22; Mismatches 67; Indels 15; Gaps 4;  
QY 10 INTTGTAGTQVSTNFTFAVRGRL-----TVLPNRVGLPINQRFILVELSNHAELSV 61  
Db 6 IKFSTEGATQSQYKQFTEALRELRGLGIIHDIPVLRDPTTVEENRRTYIVELSNSERESI 65  
QY 62 TLALDVTNAVVGVRAGNSAYFFHPDQDAEALTHLFTDVQNRYYTFAFGNGYDRLEOLA 121  
Db 66 EGVGIDVTNAVVAIRAGSQSYEL--RDAPASATYLTGTQY-RYSLRFDGSGYGLERWA 121  
QY 122 GNRLENIELGNGLPEAEISALYYVSTGCTQLPTLARSFICIQMISAEARQYIEGENWRT 181  
Db 122 HQTREQISLGLQALTHAIS---FLRSGASNDDEKARTLIIIVOMASEAARYRYSNRGVG 178  
QY 182 RIR 184  
Db 179 SIR 181  
RESULT 14  
Q945S2 PRELIMINARY; PRT; 563 AA.  
AC Q945S2  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2003 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA  
DE N-glycosidase).  
GN ABL.  
OS *Sambucus nigra* (European elder).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.

```

 7 YPIINFNTAGATVOSYTNFIRAVR-----GRLTVLPNRVGLPIINORFTLVLNSH 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 28 YPSVSENLGGAKSATYRDFLSNLKKTVATGYEWGLPVLRRESEFVQKSRFVLVPLTNY 87
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 57 AELSVTIALDVNAYVGYGAGNSATFFHPDNOEDAEI--THLFTDVQNRVYTFAGGNY 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
 88 NGNVTILAVDVNLVYVAFSGNANSYFF-----KDATEVQKSNLFTGKTQN-TLSETGNY 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
115 DRLEQLAGNLTRENIELGNGPLEAIGALYYVSTGGTQLPTLARSFICIMISEAARFOY 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
142 DNLETAANTRRSEIPELGSPLDGAITSLYHGD-----SVARSLLVLIQWYSEARFXY 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 IEGEMRTIIR 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
195 IEQEVRRSLQ 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: February 10, 2004, 16:26:37  
Job time : 25:2187 secs

Search completed: February 10, 2004, 16:26:37  
Job time : 25.2187 secs

[illegible]

OR N-PSDB; AAQ27876.

```

XX PT Prepn. of soluble recombinant polypeptide(s), esp. ricin A - by
XX PT adjusting pH and/or temp. during cultivation to increase yield of
XX PS soluble prep.
XX PS Disclosure; Fig 9; 49pp; English.
XX CC Prepn. of ricin A comprises cultivating a host, including a DNA
XX CC sequence which encodes ricin A, e.g. from pCil102, in a nutrient
XX CC medium for an initial period at a first pH value which favours growth
XX CC of the host; and cultivation the host for a further period at a pH
XX CC lower than the first pH value, and opt. cooling the host during the
XX CC terminal portion of the cultivation and harvesting the host during
XX CC the terminal portion. By adjusting the pH and temp. during the
XX CC cultivation of hosts, high yields of soluble recombinant ricin A
XX CC are obtained.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 267 AA;

Query Match 100.0%; Score 1019; DB 13; Length 267;
Best Local Similarity 100.0%; Pred. No. 5.8e-100;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60
Db 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60

Qy 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120
Db 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120

Qy 121 GNYDRLEQLAGNLENELNGPLEEASALYYSTGCTGTLPTLARSFIICQMISEAAR 180
Db 121 GNYDRLEQLAGNLENELNGPLEEASALYYSTGCTGTLPTLARSFIICQMISEAAR 180

Qy 181 FOYIEGEMTRIRYNRRS 198
Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 2
AAB19265
ID AAB19265 standard; protein; 267 AA.
XX AC AAB19265;
XX DT 19-FEB-2001 (first entry)
XX DE Amino acid sequence of a human ricin toxin A chain (RTA).
XX KW Immunotoxin; cytokine; vascular leak syndrome; VLS; lymphoma; myeloma;
XX KW Graft versus host disease; metastatic lesion tumour; tumour; immunotoxin;
XX KW ricin toxin A chain.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX Misc-difference 74..76
FT FT /note= "vascular leak syndrome (VLS) inducing motif"
XX PN WO200058456-A2.
XX PD 05-OCT-2000.
XX PF 30-MAR-2000; 2000WO-US08600.
XX PR 30-MAR-1999; 99US-0126826.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX PI Vitetta ES, Ghetie VF, Baluna RG, Smallshaw J;

```

```

XX WPI; 2000-664922/64.
XX Modifying the ability of a proteinaceous composition to induce a toxic
XX effect for reducing vascular leak syndrome, comprises identifying at
XX least one specified amino acid sequence and altering it -
XX Example 1; Page 119-120; 125pp; English.
XX The specification describes a method for producing immunotoxins and
XX cytokines with a reduced ability to promote vascular leak syndrome
XX (VLS). The immunotoxins are useful for treating graft versus host
XX disease, non-Hodgkin's and Hodgkin's lymphoma, myeloma, metastatic
XX lesion tumours and some type of solid tumours. The present sequence
XX represents an immunotoxin (ricin toxin A chain) which induces VLS.
XX The VLS-inducing motif can be mutated or deleted so that VLS is not
XX induced.
XX SQ Sequence 267 AA;

Query Match 100.0%; Score 1019; DB 21; Length 267;
Best Local Similarity 100.0%; Pred. No. 5.8e-100;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60
Db 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60

Qy 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120
Db 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120

Qy 121 GNYDRLEQLAGNLENELNGPLEEASALYYSTGCTGTLPTLARSFIICQMISEAAR 180
Db 121 GNYDRLEQLAGNLENELNGPLEEASALYYSTGCTGTLPTLARSFIICQMISEAAR 180

Qy 181 FOYIEGEMTRIRYNRRS 198
Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 3
AAR37290
ID AAR37290 standard; protein; 267 AA.
XX AC AAR37290;
XX DT 25-MAR-2003 (updated)
XX DT 09-JAN-2003 (updated)
XX DT 13-SEP-1993 (first entry)
XX DE Ricin A chain.
XX KW Type II ribosome-inactivating protein; type II RIP; gelonin;
XX KW momordin; immunoconjugate; autoimmune disease; cell killing; toxin.
XX OS Unidentified.
XX PN WO9309130-A1.
XX PD 13-MAY-1993.
XX PF 04-NOV-1992; 92WO-US09487.
XX PR 04-NOV-1991; 91US-0787567.
XX PR 19-JUN-1992; 92US-0901707.
XX PA (XOMA ) XOMA CORP.
XX PI Bernhard SL, Better MD, Carroll SF, Lane JA, Lei SP;
XX WPI; 1993-167617/20.

```



PT Analogues of type I ribosome inactivating protein - useful as  
 PT cytotoxic agents, immuno toxins for treating auto immune diseases,  
 PT cancer, graft versus host disease and selective cell killing in-vivo  
 XX  
 PS Claim 1; Page 92; 163pp; English.

XX The invention covers analogues of Type I RIPs. Ricin is a Type II  
 CC RIP whose A chain is homologous to plant type I RIPs. The analogues  
 CC of the invention have a cysteine available for intermolecular  
 CC disulphide bonding at an amino acid position corresp. to a position  
 CC not naturally available for bonding; the cys residue is located in  
 CC the C-terminal region of the analogue between a position corresp. to  
 CC amino acid 251 and the C-terminus of ricin A chain. The analogues are  
 CC pref. joined via a disulphide linkage to a molecule which specifically  
 CC binds to a target cell, e.g. an antibody fragment.  
 CC (Updated on 09-JAN-2003 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 267 AA;  
 Query Match 99.1%; Score 1010; DB 14; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-99;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62  
 Db 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62  
 QY 63 SNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTTFAGGN 122  
 Db 63 SNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTTFAGGN 122  
 QY 123 YDRLEQLAGNLRNLELNGPLPEEALISALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182  
 Db 123 YDRLEQLAGNLRNLELNGPLPEEALISALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182  
 QY 183 YIEGEMRTRIRYNRRS 198  
 Db 183 YIEGEMRTRIRYNRRS 198

RESULT 4  
 AAR63902  
 ID AAR63902 standard; protein; 267 AA.  
 AC AAR63902;  
 XX

DT 25-MAR-2003 (updated)  
 DT 27-JUL-1995 (first entry)  
 XX

DE Ricin A-chain (RTA).

KW Ricin A chain; RTA; ribosome-inactivating proteins; RIPs;  
 KW cytotoxic therapeutic agents; autoimmune disease; cancer;  
 KW graft-versus-host disease.

OS Ricinus communis.

XX W09426910-A1.

PN 24-NOV-1994.

XX 12-MAY-1994; 94WO-US05348.

XX 12-MAY-1993; 93US-0064691.

PA (XOMA ) XOMA CORP.

XX Better MD, Carroll SS, Studnicka GM, Carroll SF;  
 PI WPI; 1995-006804/01.

XX Polynucleotide(s) encoding type I ribosome-inactivating proteins

PT

PT - which are suitable for use as components of cytotoxic  
 PT therapeutic agents.

XX Example 3; Fig 1; 221pp; English.

XX AAR63902 is the ricin A chain gene product, it is analogous to the  
 CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.  
 CC RIPs are the key components of cytotoxic therapeutic agents (CTAs),  
 CC which include gene fusion products and immunoconjugates. CTAs may  
 CC be used to selectively eliminate any cell type to which a RIP  
 CC component is targeted, by the specific binding capacity of the  
 CC second component of the agent. They can be used in the treatment  
 CC of diseases where the elimination of a particular cell type is  
 CC desired, such as autoimmune disease, cancer and graft-versus-host  
 CC disease.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 267 AA;

Query Match 99.1%; Score 1010; DB 16; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-99;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62  
 Db 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62  
 QY 63 SNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTTFAGGN 122  
 Db 63 SNHAELSVTALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTTFAGGN 122  
 QY 123 YDRLEQLAGNLRNLELNGPLPEEALISALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182  
 Db 123 YDRLEQLAGNLRNLELNGPLPEEALISALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182  
 QY 183 YIEGEMRTRIRYNRRS 198  
 Db 183 YIEGEMRTRIRYNRRS 198

RESULT 5  
 AAW25136  
 ID AAW25136 standard; Protein; 290 AA.  
 AC AAW25136;  
 XX

DT 25-MAR-2003 (updated)  
 DT 02-DEC-1997 (first entry)  
 XX

DE Ricin A-chain ribosome inhibitory protein inactive precursor.

KW Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;  
 KW internal linker; Barley Translation Inhibitor; Trichosanthin;  
 KW Ricin A-chain; Abrin-A A-chain; Saporin; ST-J; Luffin A; WAP;  
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;  
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;  
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;  
 KW human immunodeficiency virus; acquired immune deficiency syndrome.

OS Synthetic.

XX US5646026-A.

XX 08-JUL-1997.

XX 07-JUN-1995; 95US-0485286.

XX 09-DEC-1992; 92US-0987927.

XX 11-JUN-1990; 90US-0535636.

XX 26-JAN-1995; 95US-0378761.

XX 07-JUN-1995; 95US-0485286.

XX (DOWC ) DOWELANCO.

XX PI Hey TD, Morgan AER, Walsh TA;  
 XX DR WPI; 1997-362934/33.  
 XX  
 XX DNA encoding pro-ribosome inactivating proteins - inactive  
 XX PT precursors of ribosome inactivating proteins; can be expressed in  
 XX PT eukaryotic cells without causing cell death  
 XX  
 XX Claim 4; Column 91-94; 186pp; English.  
 XX  
 XX AAW25136 represents a Ricin A-chain ribosome inhibitory protein (RIP)  
 XX CC which was engineered to contain a selectively removable internal peptide  
 XX CC linker sequence separating the alpha and beta units of the RIP. When  
 XX CC separated the two units regain activity and are capable of inactivating  
 XX CC eukaryotic ribosomes and hence preventing protein production. Many  
 XX CC different RIPs may be produced with an internal linker including  
 XX CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and  
 XX CC Saporin. The RIPs can be used in the construction of therapeutic  
 XX CC toxins targeted to specific cells such as tumour cells via the  
 XX CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.  
 XX CC A further use is in HIV therapy (see US4869903). There is interest  
 XX CC in expressing RIP recombinantly in host eukaryotic cells, because of  
 XX CC the capacity to provide correct post-translational processing. However,  
 XX CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting  
 XX CC in cell death. Since the inactive RIP proteins are not cytotoxic to  
 XX CC eukaryotic cells, they can be recombinantly expressed in such cells and  
 XX CC then converted to active RIP proteins.  
 XX CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 290 AA;  
 Query Match 99.1%; Score 1010; DB 18; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-99;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREHEIPVLPNRVGLPINQRFILVEL 62  
 Db 27 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREHEIPVLPNRVGLPINQRFILVEL 86  
 QY 63 SNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVYTFAGGN 122  
 Db 87 SNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVYTFAGGN 146  
 QY 123 YDRLEQLAGNLRENIELNGPGLPEEAISALYYSTGGTOLPTLARSFIICQMISEAARFQ 182  
 Db 147 YDRLEQLAGNLRENIELNGPGLPEEAISALYYSTGGTOLPTLARSFIICQMISEAARFQ 206  
 QY 183 YIEGEMTRIRYNRRS 198  
 Db 207 YIEGEMTRIRYNRRS 222  
 RESULT 6  
 AAW21699  
 XX ID AAW21699 standard; Protein; 290 AA.  
 XX AC  
 XX AC  
 XX AC  
 XX DT 25-MAR-2003 (updated)  
 XX DT 26-SEP-1997 (first entry)  
 XX DE Ricin A-chain RIP.  
 XX KW pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;  
 XX KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;  
 XX KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;  
 XX KW rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.  
 XX OS Ricinus communis.  
 XX FH Key Location/Qualifiers  
 XX FT Region 152..162

FT Hey TD, Morgan AER, Walsh TA;  
 XX WPI; 1997-309831/28.  
 XX  
 XX Inactive precursor of maize ribosome-inactivating protein - also  
 XX PT chimeric ribosome-inactivating protein precursors containing  
 XX PT internal linker sequences  
 XX  
 XX Claim 2; Column 91-94; 121pp; English.  
 XX  
 XX The sequences given in AAW21698-710 represent Ribosome Inactivating  
 XX CC Proteins (RIP's), which may be used in the construction of the  
 XX CC proRIP of the invention. The proRIP has a selectively removable,  
 XX CC internal peptide linker. The precursor sequence is incapable of  
 XX CC inactivating eukaryotic ribosomes, but can be converted by removal  
 XX CC of the linker into a protein having alpha and beta fragments and being  
 XX CC capable of inactivating eukaryotic ribosomes. RIPs are potent  
 XX CC inhibitors of eukaryotic protein synthesis. They possess a highly  
 XX CC specific N-glycosidase activity which cleaves the glycosidic bond of  
 XX CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit  
 XX CC cellular proliferation of cells, e.g. cancer cells and HIV-infected T  
 XX CC cells. The inactive proRIP proteins make it possible to provide protein  
 XX CC synthesis inhibitors with uses in practical and improved ways not before  
 XX CC possible. The RIP can be used to make cytotoxic conjugates.  
 XX CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 290 AA;  
 Query Match 99.1%; Score 1010; DB 18; Length 290;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-99;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREHEIPVLPNRVGLPINQRFILVEL 62  
 Db 27 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREHEIPVLPNRVGLPINQRFILVEL 86  
 QY 63 SNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVYTFAGGN 122  
 Db 87 SNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVYTFAGGN 146  
 QY 123 YDRLEQLAGNLRENIELNGPGLPEEAISALYYSTGGTOLPTLARSFIICQMISEAARFQ 182  
 Db 147 YDRLEQLAGNLRENIELNGPGLPEEAISALYYSTGGTOLPTLARSFIICQMISEAARFQ 206  
 QY 183 YIEGEMTRIRYNRRS 198  
 Db 207 YIEGEMTRIRYNRRS 222  
 RESULT 7  
 AAW70097  
 XX ID AAW70097 standard; protein; 332 AA.  
 XX AC  
 XX AC  
 XX AC  
 XX DT 09-APR-1991 (first entry)  
 XX DE Ricin A.

/note= "Position of possible insertion of internal  
 peptide linker sequence"

```

XX Ricin A; Met-aminopeptidase.
XX KW
XX OS Escherichia coli.
XX PN
XX EP219237-A.
XX PD
XX 22-APR-1987.
XX PF
XX 19-SEP-1986; 86EP-0307242.
XX XX
XX 06-MAY-1986; 86US-0860330.
XX PR
XX 20-SEP-1985; 85US-0778414.
XX XX
XX (CETU ) CETUS CORP.
XX PA
XX PI
XX Benbassat A, Bauer KA, Chang S, Chang SY;
XX WPI; 1987-110172/16.
XX DR
XX N-PSDB; AAN70152.
XX XX
XX N-terminal methionine free proteins prodn. - by using host
XX PT transformed with vector to express a methionine-amino-peptidase
XX PT
XX Disclosure; Fig. 4; 20pp; English.
XX PS
XX Ricin A may be produced in a form which lacks an N-terminal Met
XX CC using Met-aminopeptidase from E.coli.
XX CC
XX Sequence 332 AA;

Query Match          99.1%; Score 1010; DB 8; Length 332;
Best Local Similarity 100.0%; Pred. No. 7e-99;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGVPINQRFILVEL 62
DQ 38 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGVPINQRFILVEL 97
QY 63 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 122
DQ 98 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 157
QY 123 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQPLTLARSFIIQIMISEARFQ 182
DQ 158 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQPLTLARSFIIQIMISEARFQ 217
QY 183 YIEGEMTRIRYNRRS 198
DQ 218 YIEGEMTRIRYNRRS 233

RESULT 8
AAP70838
ID AAP70838 standard; protein; 332 AA.
XX AC
XX AAP70838;
XX AC
XX 25-MAR-2003 (updated)
XX DT 18-FEB-1991 (first entry)
XX XX
XX Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)
XX DE A protein encoded by PRA123.
XX DE
XX Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
XX KW plant toxin.
XX KW
XX Ricinus communis.
XX OS
XX FH Key Location/Qualifiers
XX FT Region 1..332
XX FT /note="leader"
XX FT Region 33..302

```

FT Region /note="A-chain"  
315..332  
FT /note="B-chain"

PN EP237676-A.

PD 23-SEP-1987.

PF 13-NOV-1986; 86EP-0308877.

PR 07-MAR-1986; 86US-0837583.

PA (CETU ) CETUS CORP.

PA (CHIR ) CHIRON CORP.

PI Piatak M;

WPI; 1987-265177/38.

DR N-PSDB; AAN70519.

XX New non-glycosylated ricin precursor and toxin etc. - are prepd.

PT by recombinant DNA procedures with specific isolation steps for

PT purer and soluble prods.

PS Disclosure; Fig 1; 112pp; English.

XX The full-length sequences encoding ricin A (AAN70520), ricin D

CC (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor

CC form were obtd. using messenger RNA to obtain a cDNA library, and

CC then probing the library to retrieve the desired cDNA inserts. The

CC library was probed using the 35-mer given in AAN70514. Figure 4 (see

CC AAN70520, AAN70522) shows the nucleotide sequences of three

CC plasmids contg. cDNA inserts obtd. by probing a cDNA library for

CC sequences encoding ricin B using the probe in AAN70517. The cDNA

CC inserts can be placed into expression vectors. Site-directed

CC mutagenesis may be used to place an ATG start codon and a HindIII

CC site at the beginning of the mature protein (see AAN70518). The

CC coding sequences of the inserts can be ligated into expression

CC vectors contg. the phoA promoter-operator and leader sequence

CC (AAN70523) and suitable retroregulators.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 332 AA;

Query Match 99.1%; Score 1010; DB 8; Length 332;

Best Local Similarity 100.0%; Pred. No. 7e-99;

Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGVPINQRFILVEL 62

DQ 38 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGVPINQRFILVEL 97

QY 63 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 122

DQ 98 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 157

QY 123 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQPLTLARSFIIQIMISEARFQ 182

DQ 158 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQPLTLARSFIIQIMISEARFQ 217

QY 183 YIEGEMTRIRYNRRS 198

DQ 218 YIEGEMTRIRYNRRS 233

RESULT 9

AAP95639

ID AAP95639 standard; protein; 332 AA.

XX AAP95639;

XX AAP95639;

DT 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)

DT 13-AUG-1990 (first entry)  
 XX Ricin A encoded by insert from plasmid pRA123.  
 DE Plasmid pRA123; ricin-A; ricin-B; cytotoxicity.  
 XX Ricinus communis.  
 KW Synthetic.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..35  
 FT /label= leader sequence  
 FT Peptide 36..302  
 FT /label=A-chain  
 FT Peptide 303..314  
 FT /label=linker  
 FT Peptide 315..332  
 FT /label=B-chain  
 XX EP335476-A.  
 XX  
 PD 04-OCT-1989.  
 XX  
 XX 19-JAN-1989; 89EP-0201162.  
 XX 08-FEB-1984; 84US-0578115.  
 PR 08-FEB-1984; 84US-0578121.  
 PR 09-FEB-1984; 84US-0578122.  
 PR 07-SEP-1984; 84US-0648759.  
 PR 20-SEP-1984; 84US-0653515.  
 XX  
 PA (CETU) CETUS CORPORATION.  
 XX  
 XX Gelfand D, Lawyer FC, Horn G, Greenfield L, Nitecki D, Kaplan D;  
 PI Platak MJ;  
 EI  
 XX WPI; 1989-286959/40.  
 DR N-PSDB; AAN91281.  
 XX  
 PT Recombinant vectors expressing ricin chains or diphtheria toxin -used for  
 PT prodn. of new immunotoxin conjugates with monoclonal antibodies, having  
 PT high cell specificity and good extracellular stability.  
 XX  
 XX Disclosure; Fig 14; 54pp; English.  
 XX  
 XX Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for  
 CC ricin A, as well as codons for 12 AAs joining the A to the B chain.  
 CC Following modification for ease of manipulation the plasmid was used to  
 CC construct expression vectors which express the conjugates in  
 CC host cells.  
 CC (Updated on 31-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 332 AA;  
 Query Match 99.1%; Score 1010; DB 10; Length 332;  
 Best Local Similarity 100.0%; Pred. No. 7e-99;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PKQYPIINFATTAGATVQSYNFIKAVRGRLLTGADVREHPVLPNVRVGLPINQRFILVEL 62  
 DB 38 PKQYPIINFATTAGATVQSYNFIKAVRGRLLTGADVREHPVLPNVRVGLPINQRFILVEL 97  
 QY 63 SNHAELSVTALDVNTAYVYGVYAGNSAYFPHPDQDAEAITHLFTDVQNRYYTFAFGGN 122  
 DB 98 SNHAELSVTALDVNTAYVYGVYAGNSAYFPHPDQDAEAITHLFTDVQNRYYTFAFGGN 157  
 QY 123 YDRLEQAGNLRENIEFGNGPLEAISAALYYSTGGTQLPTLARSFFICQMISEAARFQ 182  
 DB 158 YDRLEQAGNLRENIEFGNGPLEAISAALYYSTGGTQLPTLARSFFICQMISEAARFQ 217

QY 183 YIEGEMTRIRYARRS 198  
 DB 218 YIEGEMTRIRYARRS 233  
 RESULT 10  
 AAR70827  
 ID AAR70827 standard; Protein; 554 AA.  
 XX  
 XX AAR70827;  
 AC  
 XX 25-MAR-2003 (updated)  
 DT 31-AUG-1995 (first entry)  
 XX  
 XX Anti-cataract immunotoxin.  
 XX  
 KW Immunotoxin; heavy chain; light chain; variable region; antibody;  
 KW ricin-A; cytostatic; cataract; lens opacification; epithelial cell;  
 KW pHB19; 4197X; monoclonal antibody; MAB.  
 XX  
 OS Synthetic.  
 XX  
 XX Location/Qualifiers  
 FH Key  
 FT Peptide 1..27  
 FT /label= Sig peptide  
 FT /note= "phoA signal sequence"  
 FT Domain 28..145  
 FT /label= HEAVY  
 FT /note= "MAB 4197X heavy chain"  
 FT Peptide 148..166  
 FT /label= LINKER  
 FT Domain 169..274  
 FT /label= LIGHT  
 FT /note= "MAB 419X light chain"  
 FT Domain 276..544  
 FT /label= RICIN-A  
 FT Peptide 549..554  
 FT /label= TAG  
 FT /note= "hexa-histidine tail"  
 XX  
 XX WO9503828-A1.  
 XX  
 PD 09-FEB-1995.  
 XX  
 XX 15-JUL-1994; 94WO-US07919.  
 XX  
 XX 02-AUG-1993; 93US-0101329.  
 XX  
 XX (HOUS-) HOUSTON BIOTECHNOLOGY INC.  
 XX  
 XX Gould RM, Kelleher PJ, Wallace TL, Wood MS;  
 XX WPI; 1995-082036/11.  
 DR N-PSDB; AAQ85386.  
 XX  
 PT New single chain immuno:toxin - binds specifically to epithelial  
 PT cells, for inhibiting development of sec. cataracts after  
 PT extra:capular cataract extraction.  
 XX  
 XX Disclosure; Fig.4; 68pp; English.  
 XX  
 CC The immunotoxin given in AAR70827 comprises the heavy and light chain  
 CC variable regions of anti-lens epithelium IGG3 MAB 4197X linked to  
 CC ricin-A and a hexa-histidine tag. The DNA construct encoding the  
 CC immunotoxin was expressed from pHB19 in E. coli.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 554 AA;  
 Query Match 99.1%; Score 1010; DB 16; Length 554;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 62  
 Db 280 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 339  
 QY 63 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOEDAEATHLFTDQNRVTFAFGN 122  
 Db 340 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOEDAEATHLFTDQNRVTFAFGN 399  
 QY 123 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPTLARSFIICQMISEARFQ 182  
 Db 400 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPTLARSFIICQMISEARFQ 459  
 QY 183 YIEGEMTRIRYNRRS 198  
 Db 460 YIEGEMTRIRYNRRS 475

## RESULT 11

AAP90079  
 ID AAP90079 standard; protein; 562 AA.  
 XX  
 AC  
 XX  
 AC  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Ricin D.  
 XX  
 KW Ricin D; Ricinus communis; castor beans; Zanicariensis variety;  
 KW modified; lectin binding removed; reduced cell binding  
 XX  
 OS Ricinus communis (castor beans).  
 XX  
 PN WO8904839-A.  
 XX  
 PD 01-JUN-1989.  
 XX  
 PF 23-NOV-1988; 88WO-US04238.  
 XX  
 PR 24-NOV-1987; 87US-0124735.  
 XX  
 PA (GENY) GENETICS INST INC.  
 XX  
 PI Brown EL, Jones S;  
 XX  
 DR WPI; 1989-178366/24.  
 XX  
 DR N-PSDB; AAN90068.  
 XX  
 PT Modified ricin molecules and toxin conjugates  
 PT - in which the lectin binding function of the B chain  
 PT is removed or diminished to reduce cell binding.  
 XX  
 PS Disclosure; fig 1; 51pp; English.  
 XX  
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment  
 CC of DNA from Ricinus communis, Zanicariensis variety. Patent  
 CC discloses many modifications of ricin in which the lectin binding  
 CC function of the B chain is diminished or removed, and conjugation  
 CC to toxins to eliminate cell binding.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 562 AA;

Query Match 99.1%; Score 1010; DB 10; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 62  
 Db 38 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 97  
 QY 63 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOEDAEATHLFTDQNRVTFAFGN 122

Db 98 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOEDAEATHLFTDQNRVTFAFGN 157  
 QY 123 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPTLARSFIICQMISEARFQ 182  
 Db 158 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPTLARSFIICQMISEARFQ 217  
 QY 183 YIEGEMTRIRYNRRS 198  
 Db 218 YIEGEMTRIRYNRRS 233

## RESULT 12

AAG78304  
 ID AAG78304 standard; Protein; 565 AA.  
 XX  
 AC AAG78304;  
 XX  
 DT 27-NOV-2001 (first entry)  
 XX  
 DE Modified castor bean preprorizin (SEQ ID 10).  
 XX  
 KW Castor bean plant; preprorizin; ricin; A chain; B chain;  
 KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;  
 KW retroviral infection; anti-HIV; virucide; viral protease.  
 XX  
 OS Chimeric - Ricinus communis  
 OS Chimeric - Human immunodeficiency virus type 2.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT Protein /label= Signal peptide  
 FT /label= 25..565  
 FT /label= Prorizin  
 FT /note= "Prorizin consists of the ricin A chain, a linker  
 FT peptide, and the ricin B chain. Prorizin is  
 FT proteolytically cleaved between the A chain and  
 FT the linker to yield mature ricin"  
 FT Protein 25..291  
 FT /label= Ricin A chain  
 FT /note= "N-glycosidase"  
 FT Peptide 292..303  
 FT /label= Linker peptide  
 FT Cleavage-site 296..297  
 FT /label= HIV protease\_cleavage\_site  
 FT Protein 304..565  
 FT /label= Ricin B chain  
 FT /note= "Galactose/N-acetylgalactosamine-binding lectin"  
 XX  
 WO200160393-A1.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 15-FEB-2001; 2001WO-US05282.  
 XX  
 PR 16-FEB-2000; 2000US-0182759.  
 XX  
 PA (BECH-) BECHTEL BWXT IDAHO LLC.  
 XX  
 PI Keener WK, Ward TE;  
 XX  
 DR WPI; 2001-581908/65.  
 XX  
 DR N-PSDB; AAI64145.  
 XX  
 PT Novel composition comprising toxin e.g., ricin based antiviral compound  
 PT useful for treating viral infections such as human immunodeficiency  
 PT virus infection.  
 XX  
 PS Example 1; Page 59-63; 66pp; English.  
 XX  
 CC The sequence relates to the amino acid sequence of a modified preprorizin  
 CC protein encoded by AAI64145. The invention relates to a novel toxin  
 CC (e.g. ricin) based antiviral agent which is toxic to virus-infected

CC cells, but non-toxic to uninfected cells. The invention has anti-HIV and  
 CC virucide activities. The agent is able to enter all HIV susceptible  
 CC cells, and not just cells known to act as host cells for the virus. The  
 CC antiviral agent remains inert in a cell unless the cell is infected  
 CC with the HIV virus, where the viral protease activates it. Ricin's  
 CC mechanism of action is through inactivation of cellular ribosomes and  
 CC enhancement of binding is through inactivation of cellular ribosomes and  
 CC cell surfaces, and its cellular internalisation. The invention is useful  
 CC for treating human immunodeficiency virus infection and other viral  
 CC infections, especially retroviral infections. The antiviral agent is  
 CC activated in viral particles or early-stage infected cells, killing the  
 CC cells upon infection and effectively preventing the integration of the  
 CC viral genome into the host genome thereby preventing the latency/rebound  
 CC problem.

XX Sequence 565 AA;  
 SQ Query Match 99.1%; Score 1010; DB 22; Length 565;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILVEL 62  
 DB 27 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILVEL 86  
 QY 63 SNHAELSVTLLADVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGGN 122  
 DB 87 SNHAELSVTLLADVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGGN 146  
 QY 123 YDRLEQLAGNLRENIELGNGPLEEASALYYSTGTQTLPLARSFFICIQMISEAARFQ 182  
 DB 147 YDRLEQLAGNLRENIELGNGPLEEASALYYSTGTQTLPLARSFFICIQMISEAARFQ 206  
 QY 183 YIEGEMTRIRYNRRS 198  
 DB 207 YIEGEMTRIRYNRRS 222

RESULT 13  
 AAP70326  
 ID AAP70326 standard; Protein; 576 AA.  
 XX AC AAP70326;  
 XX AC  
 XX DT 25-MAR-2003 (updated)  
 XX DT 21-MAY-1991 (first entry)  
 XX Sequence of Ricinus communis (castor bean) Ricin toxin  
 DE (RT or ricin) E precursor encoded by prt38.  
 XX Lectin; toxin protein; cytotoxic; cytostatic; castor bean;  
 XX plant toxin.

XX Ricinus communis.  
 XX Key Location/Qualifiers  
 FH Region 1..35  
 FT /note= "leader"  
 FT Region 36..302  
 FT /note= "A-chain"  
 FT Region 315..576  
 FT /note= "B-chain"  
 XX EP237676-A.  
 XX PD 23-SEP-1987.  
 XX PF 13-NOV-1986; 86EP-0308877.  
 XX PR 07-MAR-1986; 86US-0837583.  
 XX (CETU ) CETUS CORP.  
 XX (CHIR ) CHIRON CORP.

XX Piatak M;  
 XX FI  
 XX WPI; 1987-265177/38.  
 XX N-PSDB; AAN70526.  
 XX New non-glycosylated ricin precursor and toxin etc. - are prep'd.  
 XX by recombinant DNA procedures with specific isolation steps for  
 XX purer and soluble prods.  
 XX Disclosure; Fig 14(1-2); 112pp; English.  
 XX The full length sequences encoding ricin A (AAN70520), ricin D  
 CC (AAN70525) putative ricin E (AAN70526) and RCA (AAN70524) in precursor  
 CC form were obtained, using the messenger RNA to obtain a cDNA library, and  
 CC then probing the library to retrieve the desired cDNA inserts. The  
 CC library was probed using the 35-mer given in AAN70514. Figure 4 (see  
 CC AAN70520, AAN70521, AAN70522), shows the nucleotide sequences of three  
 CC plasmids containing cDNA inserts obtained by probing a cDNA library  
 CC for sequences encoding ricin B using the probe in AAN70517. The cDNA  
 CC inserts can be placed into expression vectors. Site-directed  
 CC mutagenesis may be used to place an ATG start codon and a HindIII  
 CC site at the beginning of the mature protein, (see AAN70518). The  
 CC coding sequences of the inserts can be ligated into expression  
 CC vectors containing the Phoa promoter-operator and leader sequence  
 CC (AAN70523) and suitable retroregulators.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX Sequence 576 AA;

Query Match 99.1%; Score 1010; DB 8; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-98;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILVEL 62  
 DB 38 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPINQRFILVEL 97  
 QY 63 SNHAELSVTLLADVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGGN 122  
 DB 98 SNHAELSVTLLADVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGGN 157  
 QY 123 YDRLEQLAGNLRENIELGNGPLEEASALYYSTGTQTLPLARSFFICIQMISEAARFQ 182  
 DB 158 YDRLEQLAGNLRENIELGNGPLEEASALYYSTGTQTLPLARSFFICIQMISEAARFQ 217  
 QY 183 YIEGEMTRIRYNRRS 198  
 DB 218 YIEGEMTRIRYNRRS 233

RESULT 14  
 AAW25787  
 ID AAW25787 standard; Protein; 576 AA.  
 XX AC AAW25787;  
 XX AC  
 XX DT 25-MAR-2003 (updated)  
 XX DT 27-MAR-1998 (first entry)  
 XX Castorbean ricin.  
 XX Ricin; cytotoxin; hybrid protein; cell delivery;  
 KW cell binding ligand; translocation domain; diphtheria toxin B';  
 KW interleukin-2; T-cell lymphoma; organ rejection; therapy.  
 XX Ricinus communis.  
 XX Key Location/Qualifiers  
 FH Peptide 1..35  
 FT /label= Sig\_peptide  
 FT Protein 36..302  
 FT /label= A-domain

FT Peptide 303...314  
 FT /label= Linker  
 FT Domain 315...576  
 FT /label= B-domain

FN US5668255-A.

PD 16-SEP-1997.

XX 04-AUG-1993; 93US-0102387.

XX 27-JUN-1991; 91US-0722484.

PR 07-JUN-1984; 84US-0618199.

PR 25-APR-1985; 85US-0726808.

PR 07-JUN-1985; 85US-0742554.

PR 22-DEC-1989; 89US-0456095.

PR 14-JUN-1990; 90US-0538276.

PR 04-AUG-1993; 93US-0102387.

XX (SERA-) SERAGEN INC.

PA Murphy JR;

PI WPI; 1997-470103/43.

DR N-PSDB; AAT91638.

XX New hybrid molecules for delivery of agents to cells - comprise a

CC binding domain of a cell binding ligand and a portion of a

CC translocation domain of a protein

CC Example 4; Fig 11A-B; 30pp; English.

CC This polypeptide comprises the castor bean cytotoxin, ricin.

CC DNA (see AAT91638) encoding the enzymatic A domain and a portion

CC of the A-to-B linker peptide of ricin was used to construct a

CC ricin-diphtheria toxin B'-interleukin-2 gene that was expressed in

CC E. coli. The hybrid protein can be isolated and used to treat

CC conditions involving over-production of cells bearing IL2 receptors,

CC such as certain T-cell lymphomas and organ transplant rejection

CC crises. The hybrid inactivates ribosomes in cells bearing IL2

CC receptors, resulting in cessation of protein synthesis and death of

CC target cells. Claimed hybrid proteins comprise a translocation

CC domain and a cell binding domain from e.g. a hormone, growth factor

CC or polypeptide toxin. The hybrid molecules can be used for the

CC delivery of agents (e.g. therapeutic genes, toxins, detectable

CC labels) into cells. The use of a translocation mechanism ensures

CC that the hybrid will be effective in relatively low doses, since a

CC high proportion of the substance of interest will be taken into the

CC targeted cells. The hybrid molecules can be manufactured as a

CC single hybrid recombinant protein, permitting reproducibility,

CC consistency, and the precise control of composition.

CC (Updated on 25-MAR-2003 to correct PF field.)

FT Peptide 303...314  
 FT /label= Linker  
 FT Domain 315...576  
 FT /label= B-domain

FN US5668255-A.

PD 16-SEP-1997.

XX 04-AUG-1993; 93US-0102387.

XX 27-JUN-1991; 91US-0722484.

PR 07-JUN-1984; 84US-0618199.

PR 25-APR-1985; 85US-0726808.

PR 07-JUN-1985; 85US-0742554.

PR 22-DEC-1989; 89US-0456095.

PR 14-JUN-1990; 90US-0538276.

PR 04-AUG-1993; 93US-0102387.

XX (SERA-) SERAGEN INC.

PA Murphy JR;

PI WPI; 1997-470103/43.

DR N-PSDB; AAT91638.

XX New hybrid molecules for delivery of agents to cells - comprise a

CC binding domain of a cell binding ligand and a portion of a

CC translocation domain of a protein

CC Example 4; Fig 11A-B; 30pp; English.

CC This polypeptide comprises the castor bean cytotoxin, ricin.

CC DNA (see AAT91638) encoding the enzymatic A domain and a portion

CC of the A-to-B linker peptide of ricin was used to construct a

CC ricin-diphtheria toxin B'-interleukin-2 gene that was expressed in

CC E. coli. The hybrid protein can be isolated and used to treat

CC conditions involving over-production of cells bearing IL2 receptors,

CC such as certain T-cell lymphomas and organ transplant rejection

CC crises. The hybrid inactivates ribosomes in cells bearing IL2

CC receptors, resulting in cessation of protein synthesis and death of

CC target cells. Claimed hybrid proteins comprise a translocation

CC domain and a cell binding domain from e.g. a hormone, growth factor

CC or polypeptide toxin. The hybrid molecules can be used for the

CC delivery of agents (e.g. therapeutic genes, toxins, detectable

CC labels) into cells. The use of a translocation mechanism ensures

CC that the hybrid will be effective in relatively low doses, since a

CC high proportion of the substance of interest will be taken into the

CC targeted cells. The hybrid molecules can be manufactured as a

CC single hybrid recombinant protein, permitting reproducibility,

CC consistency, and the precise control of composition.

CC (Updated on 25-MAR-2003 to correct PF field.)

Db

218 YIEGEMTRIRYNRRS 233

RESULT 15

AAY55892

ID AAY55892 standard; Protein; 576 AA.

XX AC AAY55892;

XX DT 15-FEB-2000 (first entry)

XX XX Castor bean ricin toxin.

XX Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria;  
 translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;  
 shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;  
 cytosol; therapy; Genetic deficiency disease; enzyme; co-factor; poison;  
 adipocyte; cancer; virus; infection; antibody.

XX OS Ricinus communis.

XX PN US5965406-A.

XX PD 12-OCT-1999.

XX PF 07-JUN-1995; 95US-0488246.

XX PR 04-AUG-1993; 93US-0102387.

XX PR 07-JUN-1984; 84US-0618199.

XX PR 27-JUN-1991; 91US-0722484.

XX PR 25-APR-1985; 85US-0726808.

XX PR 07-JUN-1985; 85US-0742554.

XX PR 22-DEC-1989; 89US-0456095.

XX PR 14-JUN-1990; 90US-0538276.

XX (SERA-) SERAGEN INC.

XX PI Murphy JR;

XX DR WPI; 1999-632431/54.

XX DR N-PSDB; AAZ30663.

XX Recombinant DNA molecule encoding a three part hybrid protein used in

XX the treatment of AIDS and genetic deficiency diseases -

XX Example 4; Fig 11; 31pp; English.

The invention relates to a recombinant DNA molecule encoding a hybrid protein comprising three parts: (a) the first part comprises a portion of the binding domain of a cell-binding polypeptide ligand allowing the hybrid protein to bind to an animal cell; (b) the second part comprises a portion of a translocation domain of a naturally occurring protein selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus toxin, which translocate the third part of the across the cytoplasmic membrane into the cytosol of the cell; and (c) the third part comprises a polypeptide entity to be introduced into the cell, which is non-native to the naturally occurring protein of (b). This sequence represents the Castor bean ricin toxin sequence for use in generating the hybrid of the invention. The hybrid molecule enables the direction of appropriate therapy to affected cells, allowing them to function properly and alleviate or cure the disease. The hybrid is especially used in treating genetic deficiency diseases, by delivering to affected cells an enzyme supplying the missing function, to supplementing cellular levels of a particular enzyme or a scarce precursor or cofactor, to directing toxins or other enzymes to destroy particular cells (such as adipocytes, cancer cell, or virus infected-cells), to countering viral infections, such as HIV, by introducing appropriate antibodies to viral proteins. It is also involved in the process of getting non-therapeutic substances such as detectable labels into cells.

XX Sequence 576 AA;

XX SQ

Query Match 99.1%; Score 1010; DB 18; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-98;  
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSVYTFIRAVGRGLITGADVRIEIPVLPNRVGLPINQRFILVEL 62

Db 38 PKQYPIINFTTAGATVQSVYTFIRAVGRGLITGADVRIEIPVLPNRVGLPINQRFILVEL 97

QY 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHTLFTDVQNRVYTFAGGN 122

Db 98 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHTLFTDVQNRVYTFAGGN 157

QY 123 YDRLEQLAGNLRNTELGNGPLEEASISALYYSTGCTQLPTLARSFFIICQMISEARFQ 182

Db 158 YDRLEQLAGNLRNTELGNGPLEEASISALYYSTGCTQLPTLARSFFIICQMISEARFQ 217

QY 183 YIEGEMTRIRYNRRS 198

|||||

Query Match 99.1%; Score 1010; DB 20; Length 576;  
Best Local Similarity 100.0%; Pred. No. 1.5e-98;  
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3	PKQYPIINF	TAGATVOSYTNFIRAVRGR	TTGADV	RHEIPVLPNRVGLPINQRFILVEL	62
Db	36	PKQYPIINF	TAGATVOSYTNFIRAVRGR	TTGADV	RHEIPVLPNRVGLPINQRFILVEL	97
QY	63	SNHAELS	SVTLALDVTNAYVVG	YRAGNSAYFPHPD	NQEDAEATHLFTDVQNR	122
Db	98	SNHAELS	SVTLALDVTNAYVVG	YRAGNSAYFPHPD	NQEDAEATHLFTDVQNR	157
QY	123	YDRLEQ	LAGNLRNIE	LGNGPLEE	ISALYYYSTGGTQLPTLARSFII	182
Db	158	YDRLEQ	LAGNLRNIE	LGNGPLEE	ISALYYYSTGGTQLPTLARSFII	217
QY	183	YIEGEM	TRIRYNRRS	198		
Db	218	YIEGEM	TRIRYNRRS	233		

Search completed: February 10, 2004, 16:22:27  
Job time : 33.7583 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 11.4485 Seconds  
(without alignments)  
731.761 Million cell updates/sec

Title: US-10-083-336A-7

Perfect score: 1019

Sequence: 1 MVPKQYPIINFTAGATVQS.....ARFQYIEGEMRTIRYNRES 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/protdata/1/1aa/5A COMB.pcp:\*  
2: /cgn2\_6/protdata/1/1aa/5B COMB.pcp:\*  
3: /cgn2\_6/protdata/1/1aa/6A COMB.pcp:\*  
4: /cgn2\_6/protdata/1/1aa/6B COMB.pcp:\*  
5: /cgn2\_6/protdata/1/1aa/PCUS COMB.pcp:\*  
6: /cgn2\_6/protdata/1/1aa/backfiles1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1019	100.0	267	1	US-08-218-303-16
2	1019	100.0	267	2	US-08-338-793D-61
3	1019	100.0	267	4	US-09-538-873-1
4	1010	99.1	267	1	US-07-901-707-1
5	1010	99.1	267	1	US-07-988-430-1
6	1010	99.1	267	1	US-08-425-336-1
7	1010	99.1	267	1	US-08-488-113B-1
8	1010	99.1	267	1	US-08-477-484B-1
9	1010	99.1	267	2	US-08-646-360-1
10	1010	99.1	267	3	US-08-839-765-1
11	1010	99.1	267	3	US-09-136-389-1
12	1010	99.1	267	4	US-09-610-838-1
13	1010	99.1	267	5	PCT-US92-09487-1
14	1010	99.1	268	2	US-08-356-786-8
15	1010	99.1	290	1	US-08-378-761A-27
16	1010	99.1	290	1	US-08-485-286-27
17	1010	99.1	290	6	5248606-4
18	1010	99.1	534	2	US-08-356-786-10
19	920.5	90.3	540	1	US-08-378-761A-77
20	920.5	90.3	540	1	US-08-485-286-77
21	342	33.6	247	1	US-08-488-113B-6
22	342	33.6	247	1	US-08-477-484B-6
23	342	33.6	247	2	US-08-646-360-6
24	342	33.6	247	3	US-08-839-765-6
25	342	33.6	247	3	US-09-136-389-6
26	342	33.6	247	4	US-09-610-838-6
27	342	33.6	267	1	US-08-378-761A-74

28 342 33.6 267 1 US-08-485-286-74  
29 342 33.6 289 1 US-07-923-692C-4  
30 342 33.6 289 1 US-08-184-237-4  
31 342 33.6 289 1 US-08-482-920-4  
32 342 33.6 289 3 US-08-484-341-4  
33 342 33.6 289 3 US-08-483-502-4  
34 342 33.6 289 4 US-09-726-651A-4  
35 341.5 32.3 250 1 US-08-324-301-15  
36 329.5 32.3 250 1 US-08-378-761A-71  
37 329.5 32.3 250 1 US-08-485-286-71  
38 323.5 31.7 251 4 US-09-538-873-3  
39 312.5 30.7 255 1 US-07-901-707-6  
40 312.5 30.7 255 1 US-07-988-430-6  
41 312.5 30.7 255 5 PCT-US92-09487-6  
42 312.5 30.6 248 3 US-08-902-486-7  
43 312 30.6 290 1 US-08-245-754A-2  
44 312 30.6 290 2 US-08-597-731-2  
45 312 30.6 290 2

## ALIGNMENTS

### RESULT 1

US-08-218-303-16  
; Sequence 16, Application US/08218303  
; Patent No. 5547867  
; GENERAL INFORMATION:  
; APPLICANT: Kara, Bhupendra V.  
; APPLICANT: Hockney, Robert C.  
; APPLICANT: Fittton, John E.  
; TITLE OF INVENTION: FERMENTATION PROCESS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cushman, Darby & Cushman  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20036-5601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/218,303  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/841,533  
; FILING DATE: 26-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kokulis, Paul N.  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: PNK/3893/94908/MJW  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3000  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-218-303-16

Query Match 100.0%; Score 1019; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.3e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVPKQYPIINFTAGATVQS...YTFIRAVRGLITGADVREIPLVLPNVRGLPINQRFILV 60

Db 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNRVGLPINQRFILV 60  
QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQNYTFAFG 120  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQNYTFAFG 120  
QY 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAR 180  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAR 180  
QY 181 FOYIEGEMTRIRYNRRS 198  
Db 181 FOYIEGEMTRIRYNRRS 198

## RESULT 2

US-08-338-793D-61  
; Sequence 61, Application US/08338793D  
; Patent No. 5840521  
; GENERAL INFORMATION:  
; APPLICANT: Barth, Peter Thomas  
; TITLE OF INVENTION: VECTOR  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DABBY CUSHMAN  
; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF  
; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
; COMPUTER: IBM PC/XT/AT Compatibles  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Microsoft Word or ASCII editors  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/338,793D  
; FILING DATE: 08-Nov-5840521-94  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/842,081  
; FILING DATE: 26-Feb-92  
; CLASSIFICATION: 435  
; APPLICATION NUMBER: 9104017.0  
; FILING DATE: 26-Feb-91  
; APPLICATION NUMBER: 9109188.4  
; FILING DATE: 29-Apr-91  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kokulis, Paul N.  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: DJB/9901/215431/TGW  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-861-3000  
; TELEFAX: 202-822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 267 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
US-08-338-793D-61

Query Match 100.0%; Score 1019; DB 2; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.3e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNRVGLPINQRFILV 60  
Db 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNRVGLPINQRFILV 60

QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQNYTFAFG 120  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQNYTFAFG 120  
QY 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAR 180  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAR 180  
QY 181 FOYIEGEMTRIRYNRRS 198  
Db 181 FOYIEGEMTRIRYNRRS 198

## RESULT 3

US-09-538-873-1  
; Sequence 1, Application US/09538873  
; Patent No. 6566500  
; GENERAL INFORMATION:  
; APPLICANT: VITETTA, ELLEN S.  
; APPLICANT: GHETIE, VICTOR F.  
; APPLICANT: SMALLSHAW, JOAN  
; APPLICANT: BALUNA, ROXANA G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK  
; FILE REFERENCE: UTSD:603  
; CURRENT APPLICATION NUMBER: US/09/538,873  
; CURRENT FILING DATE: 2000-03-30  
; EARLIER APPLICATION NUMBER: 60/126,826  
; EARLIER FILING DATE: 1999-03-30  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 267  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-538-873-1

Query Match 100.0%; Score 1019; DB 4; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.3e-111;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNRVGLPINQRFILV 60  
Db 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNRVGLPINQRFILV 60  
QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQNYTFAFG 120  
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHLFTDQVQNYTFAFG 120  
QY 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAR 180  
Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTLARSFIIQMISEAR 180  
QY 181 FOYIEGEMTRIRYNRRS 198  
Db 181 FOYIEGEMTRIRYNRRS 198

## RESULT 4

US-07-901-707-1  
; Sequence 1, Application US/07901707  
; Patent No. 5376546  
; GENERAL INFORMATION:  
; APPLICANT: Bernhard, Susan L.  
; APPLICANT: Better, Marc D.  
; APPLICANT: Carroll, Steve F.  
; APPLICANT: Lane, Julie A.  
; TITLE OF INVENTION: Materials Comprising and Methods of  
; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins

NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/901,707  
FILING DATE: 19920619  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5376546and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27129/30910  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-5750  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-901-707-1

Query Match 99.1%; Score 1010; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.5e-110;  
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3	PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL	62
Db	3	PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL	62
QY	63	SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN	122
Db	63	SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN	122
QY	123	YDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFICIQMISEAARFQ	182
Db	123	YDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFICIQMISEAARFQ	182
QY	183	YIEGEMTRIRYNRRS	198
Db	183	YIEGEMTRIRYNRRS	198

RESULT 5  
US-07-988-430-1  
Sequence 1, Application US/07988430  
Patent No. 5416202  
GENERAL INFORMATION:  
APPLICANT: Bernhard, Susan L.  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Lane, Julie A.  
APPLICANT: Lei, Shau-Ping  
TITLE OF INVENTION: Materials Comprising and Methods of  
Preparation and Use for Ribosome-Inactivating Proteins  
NUMBER OF SEQUENCES: 101  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,430  
FILING DATE: 19921209  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5416202and, Greta E.  
REGISTRATION NUMBER: 35302  
REFERENCE/DOCKET NUMBER: 31133  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 267 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-988-430-1

Query Match 99.1%; Score 1010; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.5e-110;  
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3	PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL	62
Db	3	PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL	62
QY	63	SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN	122
Db	63	SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN	122
QY	123	YDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFICIQMISEAARFQ	182
Db	123	YDRLEQLAGNLRNIEIENGPLEEASALYYSTGGTQLPTLARSFICIQMISEAARFQ	182
QY	183	YIEGEMTRIRYNRRS	198
Db	183	YIEGEMTRIRYNRRS	198

RESULT 6  
US-08-425-336-1  
Sequence 1, Application US/08425336  
Patent No. 5621083  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,336  
FILING DATE: 18-APR-1995  
CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,691  
FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Thomas C.

REGISTRATION NUMBER: P-36,989

REFERENCE/DOCKET NUMBER: 31394

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-425-336-1

Query Match 99.1%; Score 1010; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.5e-110;  
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
Db	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
QY	63	SNHAELSV	TALDVTN	AYVGYR	AGNSAY	FFHPD	NQDEA	EAI	THLFTD	VQNYTFAFGN 122
Db	63	SNHAELSV	TALDVTN	AYVGYR	AGNSAY	FFHPD	NQDEA	EAI	THLFTD	VQNYTFAFGN 122
QY	123	YDRLEQ	LAGN	RENIE	LGNGP	LEEA	ISAL	YVY	STG	TQPLTARS
Db	123	YDRLEQ	LAGN	RENIE	LGNGP	LEEA	ISAL	YVY	STG	TQPLTARS
QY	183	YIEGEM	TRIRYN	RNS	198					
Db	183	YIEGEM	TRIRYN	RNS	198					

RESULT 7  
US-08-488-113B-1  
Sequence 1, Application US/08488113B  
Patent No. 5744580  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor

CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,113B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-488-113B-1

Query Match 99.1%; Score 1010; DB 1; Length 267;  
Best Local Similarity 100.0%; Pred. No. 1.5e-110;  
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
Db	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
QY	63	SNHAELSV	TALDVTN	AYVGYR	AGNSAY	FFHPD	NQDEA	EAI	THLFTD	VQNYTFAFGN 122
Db	63	SNHAELSV	TALDVTN	AYVGYR	AGNSAY	FFHPD	NQDEA	EAI	THLFTD	VQNYTFAFGN 122
QY	123	YDRLEQ	LAGN	RENIE	LGNGP	LEEA	ISAL	YVY	STG	TQPLTARS
Db	123	YDRLEQ	LAGN	RENIE	LGNGP	LEEA	ISAL	YVY	STG	TQPLTARS
QY	183	YIEGEM	TRIRYN	RNS	198					
Db	183	YIEGEM	TRIRYN	RNS	198					

RESULT 8  
US-08-477-484B-1  
Sequence 1, Application US/08477484B  
Patent No. 5756699  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
APPLICANT: Carroll, Stephen F.  
APPLICANT: Studnika, Gary M.